NO. UMTA-MA-06-0048-78-3

THE DEVELOPMENT OF MEASURES OF SERVICE AVAILABILITY

Volume II: Task Technical Reports

R.D. Leis

Battelle Columbus Laboratories
505 King Avenue
Columbus OH 43201

JUNE 1978
FINAL REPORT

DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
URBAN MASS TRANSPORTATION ADMINISTRATION
Office of Technology Development and Deployment
Washington DC 20590

18.5
.A37
no.

Technical Report Documentation Page

1. Report No. UMTA-MA-06-0048-78-3	2. Government Accession No. DOT-TSC-	3. Recipient's Catalog No.
4. Title and Subtitle THE DEVELOPMENT OF MEASURES OF SERVICE AVAILABILITY, Volume II: Task Technical Reports UMTA-78-9, II		5. Report Date June 1978
		6. Performing Organization Code
7. Author(s) R.D. Leis		8. Performing Organization Report No. DOT-TSC-UMTA-78-9, II,
9. Performing Organization Name and Address Battelle Columbus Laboratories* 505 King Avenue Columbus OH 43201		10. Work Unit No. (TRAIS) UM833/R8708
		11. Contract or Grant No. DOT-TSC-1283-2
12. Sponsoring Agency Name and Address U.S. Department of Transportation Urban Mass Transportation Administration Office of Technology, Development and Deployment Washington DC 20590		13. Type of Report and Period Covered Final Report November 1976-December 1977
		14. Sponsoring Agency Code
15. Supplementary Notes *Under Contract to: U.S. Department of Transportation Transportation Systems Center Kendall Square Cambridge MA 02142		

16. Abstract

Service availability is defined as the impingement of failures on passenger perceived service. The alternate technologies and applications for Automated Guideway Transit (AGT) systems require service availability measures (SAMs) to gage the impact of alternate reliability and maintainability (R/M) options and goals. The transit industry views various forms of passenger delay potential to be the appropriate parameters of service availability. The propensity of a system to induce delays is a complex function of R/M and operational characteristics. No single measure or model exists which can be uniformly applied to different technologies or applications. A methodology is presented to compute these relationships for simple loop and/or shuttle systems. More complex systems will require computer simulation procedures.

This is the second of three volumes. Volume I is a summary of the research effort and results. Volume III is a set of application guidelines for controlling service availability of AGT systems during the planning, procurement, and operation of the system.

17. Key Words Service Availability Measures AGT Reliability and Maintainability AGT Passenger Service		18. Distribution Statement DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 194	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

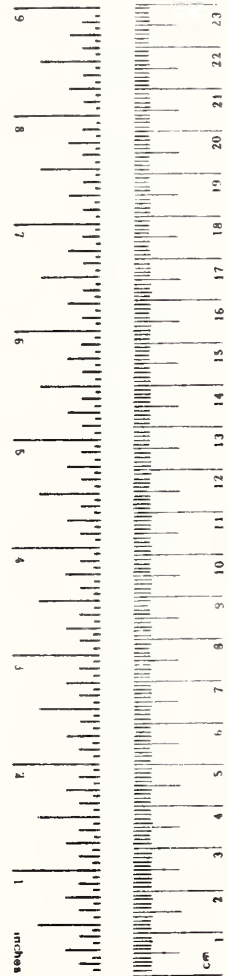
PREFACE

This three-volume set of reports constitutes the Final Report on the project "The Development of Measures of Service Availability". The project was conducted for the Transportation Systems Center (TSC) and is a part of the Urban Mass Transportation Administration's (UMTA's) "Automated Guideway Transit Technology (AFTT)" program. The objective of the project was to develop passenger-oriented measures of service availability which could be used to control the level of service provided by AGT systems throughout their life cycle.

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F

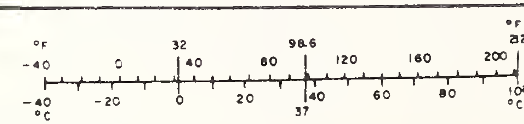


TABLE OF CONTENTS

		Page
PART 1:	Task 1 Report - Literature Review	1-i
PART 2:	Task 2 Report - Field Interviews	2-i
PART 3:	Tasks 3 and 4 Reports - Service Availability Measures Development and Guideline Document	3-i
PART 4:	Task 5 Report - Service Availability Workshop	4-i

PART 1

THE DEVELOPMENT OF MEASURES OF SERVICE AVAILABILITY

TASK 1. LITERATURE REVIEW

Contract No. DOT-TSC-1283

to

DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER

by

R. D. Leis

BATTELLE
Columbus Laboratories
505 King Avenue
Columbus OH 43201

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	1-1
2.	PRELIMINARY OBSERVATIONS REGARDING SERVICE AVAILABILITY MEASURES	1-4
3.	SERVICE AVAILABILITY IN PERSPECTIVE	1-10
4.	SERVICE AVAILABILITY CONCEPTS	1-14
	4.1 Availability Expressions of the Classical Form	1-15
	4.2 Trip Dependability Measure	1-20
	4.3 Expected Delay Measures	1-23
5.	RECENT SYSTEM SPECIFICATIONS ASSOCIATED WITH SERVICE AVAILABILITY	1-25
6.	SYSTEM PERFORMANCE MEASURES USED BY SELECTED OPERATING SYSTEMS	1-28
	6.1 Dallas-Ft. Worth Airport AIRTRANS System	1-28
	6.2 PATCO Lindenwold Line - Semi-Automatic Rapid Transit	1-29
	6.3 Bay Area Rapid Transit (BART)	1-30
	6.4 Seattle-Tacoma International Airport (Sea-Tac)	1-31
	6.5 Morgantown PRT	1-31

APPENDIX

BIBLIOGRAPHY 1-33

LIST OF FIGURES

Figure 1-1.	Sample Relationship Among Various Service Performance Measures as Applied to a Simple Hypothetical System	1-6
Figure 1-2.	Relationships Between Multiple Criteria for Specification of Availability	1-8
Figure 1-3.	Logic Tree Showing Factors Influencing Passenger Perception of Service	1-11
Figure 1-4.	Example Operating Cycle of Simple Series Transit System	1-16

LIST OF TABLES

Table 1-1. Examples of Types of Specifications 1-26

1. INTRODUCTION

For some time, there has been considerable concern over the definition, measurement, and specification of a transportation system's effectiveness for providing service to its passengers in the face of the failure characteristics and consequences inherent in the design and operation of that system. Trip reliability, schedule adherence, compactness of trip time distribution, average delay, expected delays, availability, dependability, etc.: all are used at one time or another to describe this effectiveness measure. The net result is a welter of nonstandard terminology which serves not only to confuse the analyst but also to totally mask performance comparisons among alternative systems.

Accordingly, this study (a part of UMTA's Automatic Guideway Transit Technology program) is aimed at developing a set of measures for "service availability" which will be meaningful, readily understandable, and acceptable to transit operators, suppliers, and interested Government agencies. Service availability is defined in a generic sense as a measure of the impingement of equipment failures on the operation of a transit system as perceived by the users and operators.

The first task of this study was an in-depth review of existing literature dealing directly or indirectly with the generic subject of service availability. The purpose of this document is to report the results of this task.

In this task, over 100 papers, textbooks, and sumposium proceedings were reviewed. The Appendix is a bibliography of material reviewed. In summary form, several observations are as follows:

(1) While the literature abounds in seemingly different approaches to the measurement of service availability, the generic concepts involved are few. Major differences rather appear in transit system models used and degree of mathematical complexity employed.

(2) No single measure appears to reflect all characteristics presumed to be important to passenger perception of service.

(3) Measures which approach direct correspondence with passenger perception are both difficult to compute in a design sense and measure in an operating sense. Hence, it appears that if this is desired, computer simulation techniques will be required.

(4) From our current perception of the users of a service availability measure (e.g., operators, suppliers, planners), it appears that the role of passenger perception of failure-induced delays is best dealt with in the planning phase to arrive at allowable transit system equipment performance. This latter measure would then form the basis for specifications and operational performance monitoring. Periodic measurement of actual passenger impact (e.g., through statistical techniques) would serve to verify the planning models.

(5) Most recent specifications for transit systems specify, in more or less detail, required reliability and maintainability performance. Only in a few cases have more user-oriented measures been incorporated.

(6) Most operating systems gather detailed data on equipment failure rates, repair rates, and downtimes (in and out of service). None, to our knowledge, directly measures passenger-related impacts of these failures.

These points are amplified in the sections which follow. Section 2.0 presents some observations on service availability gleaned from the literature review. Section 3.0 discusses the concept of service availability in the perspective of its role in the overall performance of a transit system. Section 4.0 presents a review of the basic concepts of service availability as discussed in the literature. Section 5.0 briefly discusses performance measures of selected existing systems as they are currently calculated. This information is derived from the papers presented at the 1976 UMTA AGT Service Availability Workshop. Section 6.0

briefly summarizes recent system specifications which impact service availability. The Appendix is a bibliography of documents reviewed in this task.

It will be noted that the text is not keyed to specific references. This is because, in most cases, the references only obliquely treat the subject and, in those cases where a full treatment is presented, they are not proposing a unique concept. In the merging of all these resources, it was rare that specific references were appropriate or warranted.

2. PRELIMINARY OBSERVATIONS REGARDING SERVICE AVAILABILITY MEASURES

As will be discussed in the following sections, the literature does not suffer from a lack of information on appropriate measures for service availability. However, as discussed in Section 4.0, what appear to be fundamentally different measures are in reality different models or procedures for computing one of three generically different measures.

Type I. Measures of the classical availability form

$$\text{Availability} = \frac{\text{Successful Time}}{\text{Total operating time}},$$

where the elements of the fraction may be expressed in terms of system hours, vehicle hours, or passenger hours.

Meaning: Likelihood of being in a successful state at any random time during use.

Type II. Measures of the classical dependability form

$$\begin{aligned}\text{Dependability} &= \text{probability of success}\\ &= \text{availability} \times \text{reliability},\end{aligned}$$

where the elements are generally computed on a per trip basis for either vehicles or passengers.

Meaning: Likelihood of not incurring a delay during a given period of use.

Type III. Measures of the expected delay form

$$\text{Expected delay} = \text{probability of delay} \times \text{avg. duration of delay},$$

where the elements are generally computed on a passenger trip or vehicle-trip basis.

Meaning: Average delay of a passenger on a typical trip.

To illustrate the differences between these types of measures, it is useful to evaluate them for a hypothetical system. The system model used is extremely simple, consisting of a link which connects a passenger between an origin and a destination. The characteristics of the link are such that any failure causes a delay to the passenger. Furthermore, his delay pattern and duration are precisely the same as those of the system.* (There is sufficient excess

* Obviously, the simple model does not represent a real system. It is used only to illustrate the types of measures and their characteristics.

capacity to ensure that queues resulting from previous failures do not introduce an additional delay event.) The failure characteristics of this system are further described by a mean-time-between-failure (MTBF) and a mean-time-to-restore normal service (MTTR). Figure 1 shows curves of constant availability, dependability, and expected delay per trip as a function of MTBR and MTTR for this system.

If this simple system were specified on the basis of a Type I measure, say availability = 0.99, the system supplier would have complete freedom to select his design point at any point along or to the right of line A in Figure 1. Obviously, where he selects his design point will have a significant influence on the resulting dependability and expected delay performance. Hence, if delay-type measures are considered important (and they are) availability measures, by themselves, are not appropriate.

If the system is specified in terms of allowable delay probability or minimum allowable dependability, say 0.957, the system supplier would have the freedom of selecting his design point along or to the right of line B in Figure 1. As can be seen, where he selects his design point will significantly influence the expected delay performance of the resulting system. Again, if this is considered important, a dependability measure by itself fails as a useful service availability measure.

If the system is specified by an allowable average delay per trip, the system supplier selects his design point on or below line C of Figure 1. If expected delay (or some variation) is really the measure of concern then obviously, these curves provide the "best" specification from the standpoint of sensitivity to passenger perception. However, as seen in Figure 1-1, a constant expected delay may be achieved with many different dependability values. If frequency of delays is important, an expected delay criterion, by itself, is not responsive.

Therefore, it would appear that, if delay frequency and duration are important passenger-perceived performance parameters to be described in a useful service availability measure, the choice lies between Types II and III measures. These measures share another characteristic. They are both difficult to compute and measure for compliance. In determining an expected

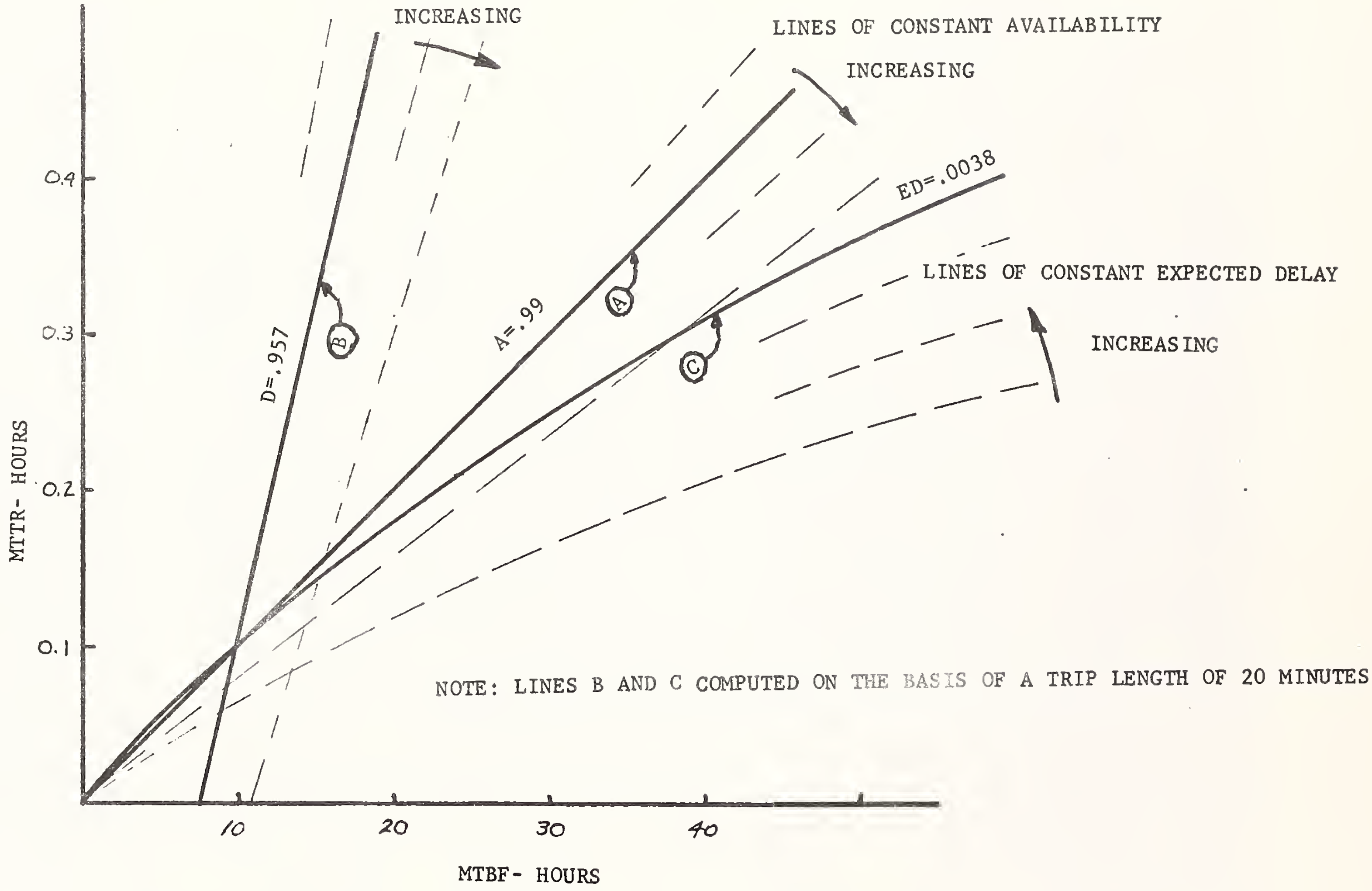
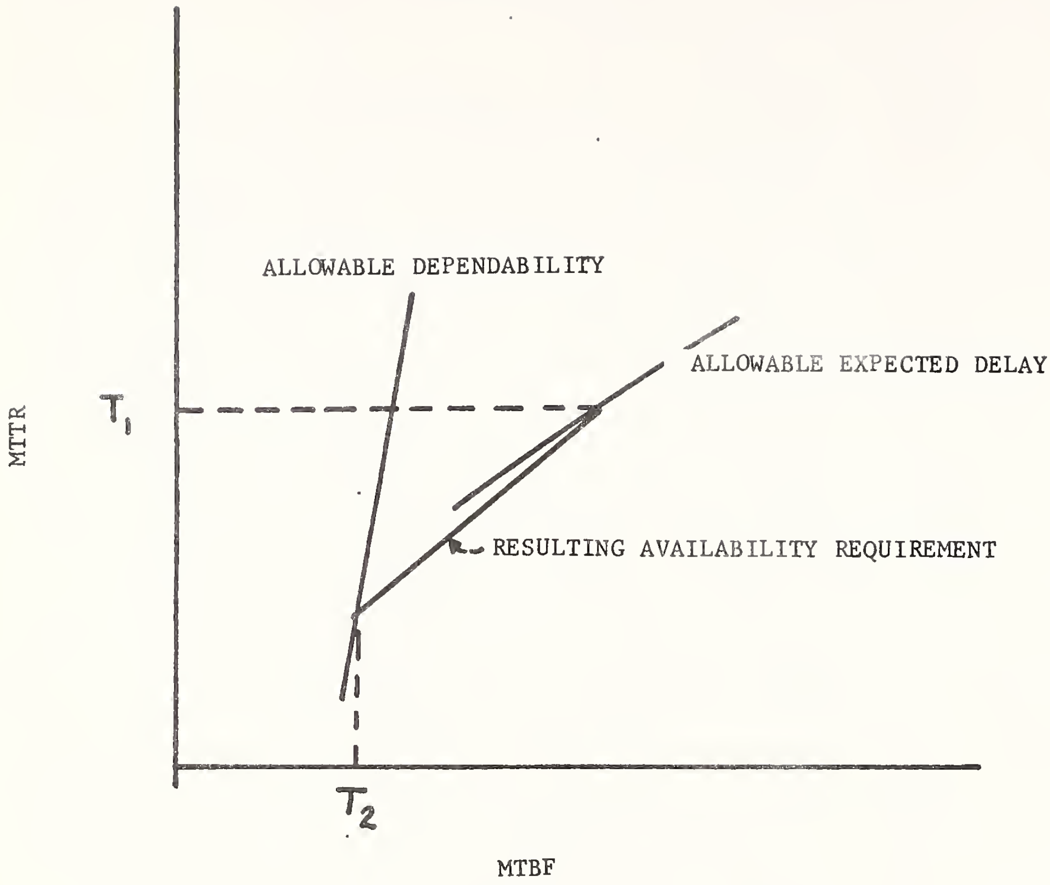


FIGURE 1-1. SAMPLE RELATIONSHIP AMONG VARIOUS SERVICE PERFORMANCE MEASURES AS APPLIED TO A SIMPLE HYPOTHETICAL SYSTEM

performance for a given transit configuration, some form of computer simulation is indicated, except for extremely simple systems. Thus, the expense of predicting performance in a design phase may be prohibitive. Similarly, compliance testing may be quite tenuous.

Therefore, perhaps some adequate proxy measure is indicated. For example, during the planning phase of a transit system, simulation could investigate the ramifications of delay frequency and duration expected for the average passenger to establish a zone of acceptable system performance. For example, referring to Figure 1-1 in the lower left-hand corner, expected delay is approximately linear and, hence, correlates with availability which suggests that availability is a reasonable proxy for expected delay, within some limited range of system MTBF and MTTR. Through simulation during the planning and preliminary design phase of a transit system, the impact of system failures on expected delay would be examined together with the implications of these failures in terms of an availability expression, for example, a vehicle uptime ratio. Within the range of acceptable expected delays, the corresponding values of vehicle availability would then be established which could be used as the system specification and performance monitoring criterion. Graphically, this is illustrated in Figure 1-2. As shown in this figure, dependability considerations could also be included to further restrict the permissible range. A specification like this would be relatively easily demonstrable in both the design and operating phase. This process suffers in one major respect, however. For two systems, which vary in degree of complexity or projected passenger demands, achieving the same service level will require different vehicle availability requirements. Therefore, comparison between systems in different locations providing different services cannot be made directly on the basis of this measure. To establish how well they are doing, in passenger perception terms, one must revert back to the original algorithm which led to the specified availability measure.

The concepts discussed here are not offered as conclusions but only as observations. It appears, based on current knowledge, that no



T_1 = MAXIMUM MTTR (From Expected Delay Considerations)

T_2 = MAXIMUM MTBF (From Dependability Considerations)

FIGURE 1-2. RELATIONSHIPS BETWEEN MULTIPLE CRITERIA FOR SPECIFICATION OF AVAILABILITY

single measure can satisfy all needs. Furthermore, as the measure approaches something that a passenger perceives, its relevance in terms of required system performance becomes lost and probably requires computer simulation to reestablish it. The relationships discussed in this section are all based on a simple model. However, while the values will be different for real systems, it is believed that similar relationships exist.

3. SERVICE AVAILABILITY IN PERSPECTIVE

As a user-perceived attribute of a transportation system, service availability, as defined in this program, probably does not exist by itself. Transit system users view the output service of a transit system as a single product, composed of a normal expected performance and perturbations about this norm--irrespective of the source of these perturbations. Hence, to gain insight into an appropriate measure for service availability, one has to take a broader view of a passenger's perception of transit system service. Beginning with basic attitude surveys(24)*, one finds that "transportation reliability" is a performance measure which ranks high in a passenger's evaluation of the acceptability and desirability of a transit system from a modal attractiveness standpoint. This attribute has not been dissected into its components nor has it been treated as a design variable in transportation planning and modal-split analyses. As a result, transportation service reliability has not been defined nor has a scale of "goodness" been determined to permit tradeoff analyses to be performed. Detailed attitude surveys have attempted to shed some light on this attribute. However, no two surveys used the same terminology in questions pertaining to service reliability, hence direct correlation of results was not possible. Examples of the components of service reliability addressed in various surveys are

Reliability of destination achievement
Waiting time
In-transit delays
Missed connections
Seat availability
Trip-time variance.

While these are different characteristics, they have sufficient commonality to at least describe the nature of the service reliability concept as the freedom from unexpected variations in scheduled performance.

These unexpected variations can arise from a number of sources--they do not all result from equipment failures. This is illustrated in Figure 1-3, where the left-hand branch coincides with the service availability definition used in this study. The lower portions of this branch describe the two major system characteristics which determine service availability.

* Numbers refer to specific references in the Bibliography.

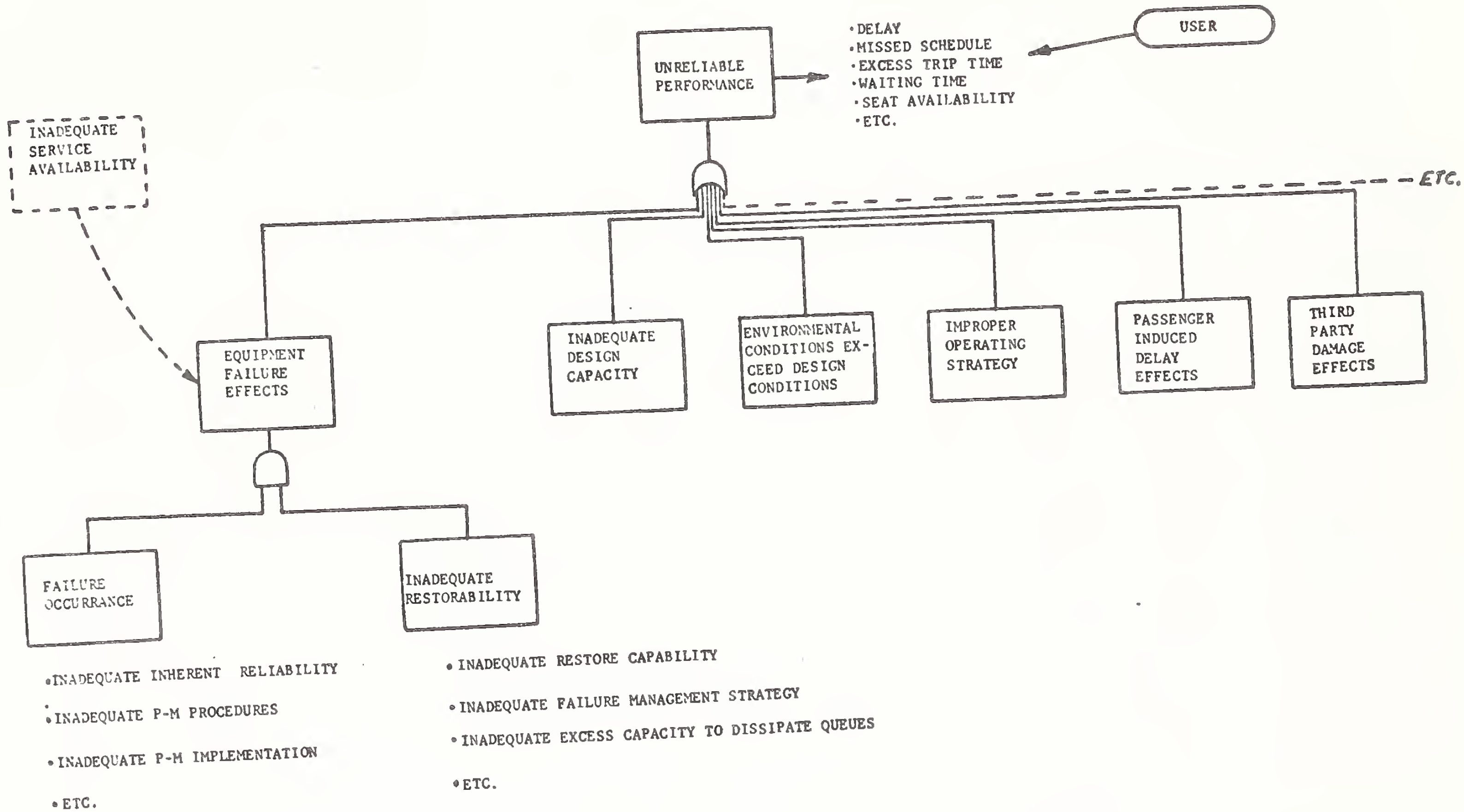


FIGURE 1-3. LOGIC TREE SHOWING FACTORS INFLUENCING PASSENGER PERCEPTION OF SERVICE

(1) The failure characteristics of the system under consideration including the type, frequency, and system impact of failures. These are controlled in the design phase under the general discipline of reliability engineering and in the operational phase through effective preventive maintenance (PM) programs.

(2) The recovery characteristics of the system; that is, the capability of the system, through design and/or operation, to recover quickly from a failure to minimize the effects of the failure in terms of user perceived performance. This is controlled in the design phase under the general discipline of maintainability engineering and in the operational phase through effective failure management procedures.

The concept of "service availability" is, therefore, that of a transfer function relating these system characteristics to system performance as viewed by the passenger (which performance, as was pointed out earlier, is not well defined). With such a simple purpose, why are there so many different measures proposed or used? There are many reasons, but three general reasons appear to dominate the answer to this question.

(1) There are various interpretations of the appropriate user perceived service parameter. While the literature has revealed almost universal consensus that some measure of <u>delay</u> is appropriate; there is considerable divergence of opinion as to the approximate transit system failure performance which characterizes its propensity to induce delay. (These form the major types of measures discussed in the next section.)

(2) The design and operating characteristics of transit systems vary greatly. Hence, even though a measure (e.g., probability of delay) might be common between two system analyses, the mathematical formulation of the availability transfer function may be quite different--giving the appearance of a different measure. In like manner, many papers may treat the same system with varying degrees of rigor, different assumptions, and different levels of mathematical complexity, again giving apparent differences in the availability concept being explored.

(3) Service availability is being addressed at various phases in the life cycle. In the operational phase, service availability is generally dealt with in terms

of observed performance and not related to failure rates, etc., as shown in Figure 3. Other studies deal with early planning activities or simply academic situations where emphasis is placed on general system level failure rates and restore options. Yet other studies deal with relatively fixed hardware concepts, with system, subsystem, and perhaps component failure data upon which system level service availability is being based. Like (2) above, the passenger measure may be the same but the models or transfer functions are quite different.

The following section describes and discusses the service availability concepts derived from the literature and will illustrate these points. Before leaving this section, however, two points must yet be made.

(1) As illustrated in Figure 1-3, service availability is but one element determining overall system performances as viewed by the passenger. "Good" service availability does not ensure a good system. While the study is aimed at service availability, evaluating the effects of other parameters indicated in Figure 3 must be recognized by the planning community.

(2) This study is directed toward the definition of service availability; i.e., the dimensions of a system which characterize its performance in the eyes of the passenger. It is quite another problem (and not within the scope of this program) to determine what the <u>level</u> of such performance should be. There is little in the literature suggesting adequate or required levels. This may well prove to be the most significant restriction to utilizing service availability to define system reliability/maintainability requirements.

4. SERVICE AVAILABILITY CONCEPTS

This section discusses the various service availability concepts discussed or promoted in the literature in what might be termed a theoretical or academic environment. Most of the thrust behind these articles is to demonstrate an assessment methodology or model for computation purposes. The measures and models are not discussed with reference to any specific application but are rather promoted as "general" purpose. These articles point out the general "multistage hypothesis" requirements of current service availability treatments, such as

(1) Hypotheses regarding passenger sensitive performance parameter

(2) Hypotheses regarding system performance characteristics which influence this parameter

(3) Hypotheses regarding system element interaction which influences these performance characteristics

(4) ....Etc.

The first hypothesis, as mentioned in the previous section, is that the passenger perceives delay as a measure of "badness" in system performance. While the precise treatment of delay might be different among authors, all subscribe to its paramount importance.

The second hypothesis is not universal among authors. There are basic differences among approaches to a system performance measure which relates to delay potential. These form the major subdivisions in this section.

The remaining hypotheses illustrate even less universality which is understandable because of the different transit system scenarios being addressed. These variations are not treated explicitly in this report but are briefly discussed within the context of the measures being evaluated.

Returning to the second hypothesis, while many different mathematical formulations exist in the literature, the service availability measures discussed can be placed in one of three general categories.

(1) Measures which are based on time and uptime or "something delivered" to "something promised". Classical availability expressions exemplify this category.

(2) Measures based on the propensity of the system being evaluated to cause delays in a random, average trip.

(3) Measures based on the magnitude of delay expected on a random or average trip.

Prior to discussing these service availability measure types, it is useful to establish a pictorial representation of system operating performance which can be referenced in each discussion to illustrate commonalities and differences. Figure 1-4 shows, in simple form, the operating cycle of a hypothetical transit system, subsystem, or component.* As illustrated, an operating cycle consists of a period of uptime or successful operating time until a failure occurs (at t_1) which renders the system (or subsystem or component) inoperable. There follows a period of time in which repair, removal, or replacement is performed to restore normal operation. Each system or subsystem will display its own operating cycle pattern depending on its failure modes, frequencies, effects, and restorability. In concept, service availability is a measure which relates these patterns to system performance parameters of concern to passengers.

Superimposed on this figure is a sample trip of time length T, commencing at T_s and ending at T_f. This simple diagram is useful in the discussions which follow.

4.1 Availability Expressions of the Classical Form

This section discusses the measures which fall in the first category above. The best known expression in this category is

$$A = \frac{\text{Uptime}}{\text{Uptime + Downtime}} \tag{1}$$

where A = Availability

Uptime = Total operating time when system is not down due to failure

Downtime = Total system downtime hours due to a failure.

---

* Like the simple model discussed in Section 2.0, many liberties were taken with this example to illustrate a point.

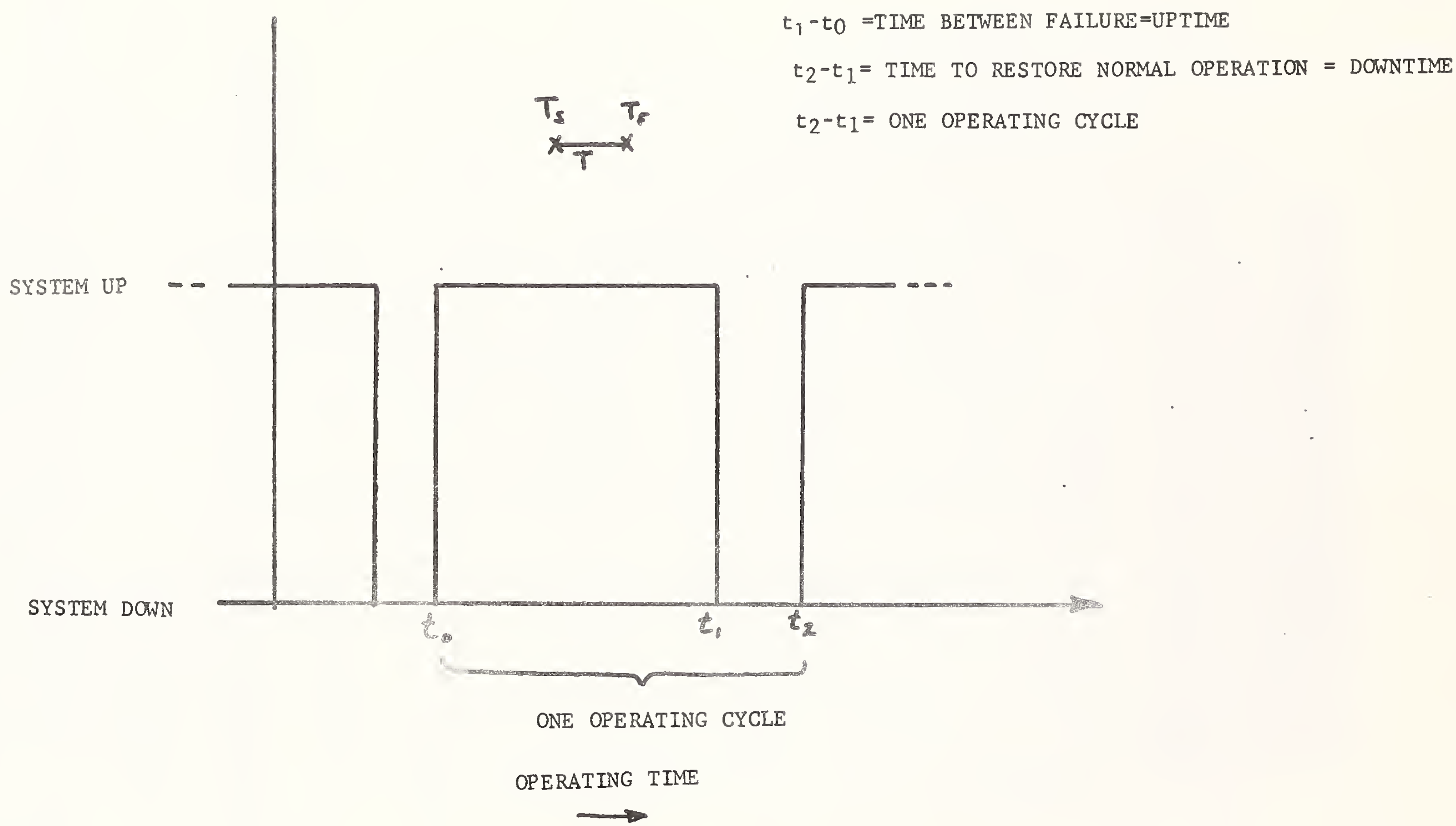


FIGURE 1-4. EXAMPLE OPERATING CYCLE OF SIMPLE SERIES TRANSIT SYSTEM

In terms of system design/operating characteristics, this measure can be related to failure rates and restore times as follows:

$$A = \frac{MTBF}{MTBF + MTTR} \tag{2}$$

where MTBF = Mean time between system failures

MTTR = Mean time to restore normal service following a failure.

In terms of Figure 1-4, this measure reflects the area under a "mean" operating cycle curve. The literal interpretation of a value for this measure of availability is "the probability that the system is in an operational state at any time within its scheduled operating period". From a passenger viewpoint, this can be interpreted as the probability that the system will be operating when he requests service. (Referring to Figure 1-4, this is the probability that T_s will fall between t_0 and t_1.)

This measure requires a minimum of data to compute in an operating environment and affords a maximum flexibility to manipulate system failure rates and restore capabilities within specified cost limits. From an operator viewpoint, it is a good summary measure of performance--reflecting a productivity or equipment utilization measure.

This measure, while incorporating the causes of delays, does not directly indicate the likelihood or severity of delays which the passenger can expect. However, its main drawback as a useful parameter in transportation performance measurement is that it characterizes a system which is either totally "up" or totally "down". One would find great difficulty in finding an existing or proposed transit system which exhibits this characteristic.

To overcome this deficiency, numerous variations exist which involve redefinition of system, uptime, downtime, and failure. For example,

$$A = \frac{\Sigma \text{ vehicle successful operating hours}}{\Sigma \text{ vehicle successful operating hours} + \Sigma \text{ vehicle hours of delay}} \tag{3}$$

Expressing availability in this form recognizes that individual vehicle delays may not render the entire system inoperable. (In a closely coupled system, where a single vehicle delay results in a total system delay, a measure computed by Equation 3 would be identical to that computed by

Equation 1.) Like Equation 1, Equation 3 measures the operational state of the system. It is easily calculated in an operational environment. In a planning environment, detailed failure mode/effect/restore calculations can be made, at varying levels of complexity to arrive at estimates of availability. Care must be taken to include vehicles delayed as a secondary effect of failures elsewhere in the system. Conceptually, one could also count as delays the excess time required to reroute vehicles around the failure if such an option exists. The mathematics of these calculations becomes extremely complicated, however, and computer simulation may be required.

This measure lends itself to quantification under an assumed allowable delay. Many authors subscribe to the theory that long delays are less tolerable than short delays, even though over a period of time the total delay experienced by the passenger may be the same for both delay types. Furthermore, there is probably a minimum delay to which a passenger is insensitive. Delays of 3 to 5 minutes or 10 percent of the normal trip time are generally proposed. Using Equation 3 under such an assumption, one would simply truncate all delays at the acceptable limit. Availability expressions, such as (3) or (1), have an inherent problem in the definition of delay. When does it start? When does it stop? Presumably this would be taken as some deviation from normal operation. However, after a certain amount of delay, one merely slips one slot in the schedule. Modeling techniques must recognize this. Delay for demand responsive systems is equally difficult to define. However, the major problem with this measure is, like (1), that it lacks sensitivity to a user expectation of delay likelihood or duration. Furthermore, it lacks sensitivity to delays to passengers which result from queuing at stations as a result of unavailability.

This latter problem can be dealt with by, again, redefinition of terms.

$$A = \frac{\Sigma \text{ uptime in passenger hours}}{\Sigma \text{ uptime in passenger hours} + \Sigma \text{ passenger hours of delay}} \qquad (4)$$

where passenger hours of delay includes station wait time beyond the normal expected value. This formulation essentially superimposes a passenger demand on a system characterized by Equation 3, allowing passenger demands to build

up on the platforms as failures in the system restrict its capacity.

This measure states that, on the average, for every hour of passenger interaction with the system, he will see (1-A) hours of delay. However, caution must be exercised in relating this value to expected delays for a finite trip or number of finite trips. The above values are based on cumulative averages and are generally insensitive to a trip duration. Availability expressions reflect ratios of uptime to downtime and not the magnitude of either. However, it is the magnitude of each with relation to the average trip time which determines both probability of delay and the average value of delay encountered. This will be discussed in later sections.

The measure expressed in Equation 4 does account for passenger wait time and queue dissipation capability (excess capacity) in the system. As such, it is closer to a passenger-based measure than (1) or (3). Evaluating this measure for a design concept requires some form of computer simulation which models the passenger movements through the system. As a measure of achieved performance in an operating system, Equation 4 requires detailed passenger movement data in terms of trip time or delay time. This requires the tracking of individual passengers through the system. In addition, an accurate algorithm for defining the normal expected values for these parameters must be available. Few, if any, systems are capable, technically or economically, of collecting these data.

For an operating system, such a measure could prove useful from the standpoint of how well it performed but offers little in the way of system management information except in the evaluation of failure management strategies and practices.

From the design specification standpoint, such a measure would first have to be stripped of the passenger loading influence to yield useful system design/operating requirements which can be manipulated in the design process. If the issue of required reserve capacity is to be left to the system suppliers, then such a measure might be appropriate for a general specification.

Because of the simulation requirements and the large number of assumptions regarding expected passenger demands, trip durations, trip origins and destinations, it would appear that a measure such as Equation 4 more properly lies in the planning domain. Along with determining the general nature of the desired system, station locations, nominal capacity requirements, and other related information, passenger response to various system failures and operating strategies could be determined to establish required levels of system performance to be used as specifications. Such specifications may involve a measure such as Equation 3, together with required reserve capacity or failure management strategies to minimize queue development under the system failure characteristics.

An extension of Equation 4 allows for weighting delay time to reflect an increasing "annoyance" with increases in delay time.

$$A = 1 - \frac{\sum_i \beta (\text{Delay Time}_i)^{\alpha}}{\text{Total Time}} \tag{5}$$

where β and α are annoyance factors. Generally, expressions like this use only one of these factors. This form is applicable to system hours, vehicle hours, or passenger hours. It has not been used in examples of real-world situations. It bears all the problems of previous measures with the addition of the need to define and evaluate the annoyance factors.

4.2 Trip Dependability Measure

The availability-type measures discussed in the previous section all relate to areas under a "goodness" curve, such as Figure 4. There is no comprehension of a trip length or starting time. As a result, two systems may exhibit the same availability measure but have different values for the duration of the "good" and "bad" periods. As such, they fail to measure what the average user of the system is likely to see in terms of delays, duration, or frequency.

Trip dependability measures attack the probability of delay directly. Essentially, these measures deal with availability in terms of the system failure and restore characteristics with the added consideration of the relationship of a random trip to these characteristics.

The probability of delay measure is the combination of two probabilities.

(1) The probability that the passenger will not experience a delay prior to starting his trip due to a failure in the system.

(2) The probability that, having started his trip, the passenger will not be delayed by a system failure.

Mathematically, the measure is computed as

$$D = A \cdot P_s \tag{6}$$

where D = Trip dependability.

A = Steady-state availability of that portion of the system which is required by the passenger or which can influence his delay probabilities. This is the probability that all required segments of the system are operational when required by the passenger.

P_s = Probability of successfully transiting the system; i.e., incurring no delay due to system failure while on the system.

Many models exist for computing trip dependability. These models differ in four major respects.

(1) They are sensitive to a particular system configuration. The "system" required for an average trip, therefore, differs.

(2) They differ in level of system subdivision employed. Hence, different relationships for computing A and P_s exist at the system level.

(3) They differ in mathematical rigor. For example, some models may treat system availability at the time the trip is initiated. Others may treat subsystem availability only at the time it is required. This allows certain downstream elements to be in a failed state initially. Their availability when required is then a function of the restore time for that particular failure. Similarly, downstream failures are permitted during the trip as long as service is restored prior to their use.

(4) Some models allow a delay, hence, failures are permissible as long as they are restored within the permissible delay time.

These models tend to be the most complicated of all used to assess service characteristics. It is not uncommon to see them spanning several pages in a report. One is necessarily impressed by the mathematical rigor employed. However, mathematical complexity does not imply applicability or accuracy of results. One is better off to apply simple mathematics to a good model than to use complicated mathematics in a poor model. In this regard, one has to be sure that his system definition is appropriate, including in his system all vehicles whose delay would impact the subject passenger's vehicle. Furthermore, one should comprehend in his availability expressions not only whether the system is up and running, but also whether the system is available in light of the queues which may have developed because of a previous failure.

As would be expected, except for simple system models and probability expressions, some type of computer simulation is required to effectively deal with dependability measures to evaluate effects of failure rates and restore times. Notwithstanding this, trip dependability is considered to be more of a passenger-related performance measure than classical availability form of measurement.

From the operator viewpoint, dependability measures are currently considered to be good measures of performance. For example, airlines use "on-time performance" as a comparative measure of service. Some transit operators use a similar measure. For example, trip dependability may be defined as follows:

$$D = \frac{\text{Number of Successful Trips}}{\text{Total Trips}} \tag{7}$$

where A successful trip may be a zero delay trip or a trip with some maximum allowable delay. Methods may also be employed to account for skipped stations and annulled trains.

A trip is generally defined as one end-to-end movement of a vehicle or train.

Because of the definition of terms, Equation 7 measures vehicle-trip dependability which does not yield a direct correlation with passenger-trip dependability. However, discounting delays due to queues, the impact of skipped stations, and the impact of annulled trains, the two measures are comparable.

One could measure passenger-trip dependability directly but the data gathering problem is identical to that for Equation 4. In fact, the data needs are identical.

From a specification standpoint, trip dependability specifications would permit the determination of allowable failure rates and restore requirements. Complicated relationships exist and computer analysis would probably be required. As was the case for Equation 4, it is perhaps more appropriate to conduct these analyses in the planning phase to determine the required system performance which can then become part of the specification. Like the classical expressions, regardless of where such simulation or determination of system requirements is performed, it is the resulting system requirement which would probably become part of the specification or promised performance of a transit system.

The major problem with dependability measures is that they do not yield insight into the duration of the delays. Two system configurations with the same propensity to induce delays in a typical trip will be viewed quite differently by passengers, depending on the relative duration of resulting delays. This gives rise to expected delay measures discussed in the next section.

4.3 Expected Delay Measures

Expected delay measures have been suggested to overcome the insensitivity to delay duration of dependability measures. Basically, expected delay measures modify dependability measures by the average duration of delay, if one occurs. In equation form

$$ED = P_d \cdot D_d, \tag{8}$$

where ED = Expected delay per trip

P_d = Probability of encountering a delay on a trip (1-dependability)

D_d = Avg. duration of delay if a delay event is encountered.

(Units may be vehicle or passenger oriented).

One could state, in a specification sense, an allowable average delay per passenger trip, or distribution of allowable delays. Working

through Equation 8, one could determine required ranges of system level failure rates and restore times. Like dependability measures, the mathematical complexity of doing this may be considerable, even using an interactive approach. Again, computer simulation would probably be required to fully examine the implications of failure rates and failure management strategies.

From an operating system viewpoint, measurement of passenger expected delay per trip would involve gathering data on total passenger delays encountered and divide by the total number of trips taken. Measurement of the vehicle-based expected delay would be easier but would require considerably more delay information than is currently collected in most operating systems.

5. RECENT SYSTEM SPECIFICATIONS ASSOCIATED WITH SERVICE AVAILABILITY

During the past few years there has been an increasing effort within the transit industry to make quality service, as viewed by the passenger, an explicit requirement of specifications and other documents describing transit system capabilities and performance. This action has been prompted by the realization that the quality of passenger service (both real and perceived) must be high, and that this can be achieved only by a specific and concerted effort at all levels of the system planning and design process.

However, this thrust has only evidenced itself, in general, with words indicating a general goal of providing service. Most specifications have been given in terms of system/subsystem allowable failure rates and mean-time-to-restore service. The thrust of these specifications, in addition to establishing performance requirements, has been to firmly establish the terminology and compliance measurement technique to be employed--a very necessary part of any specification. In most cases, presumably, it was assumed that achievement of the specified conditions would provide good service to the passenger. However, this relationship was not established. Only in a few instances was a further qualification on required performance, from a passenger viewpoint, treated in any quantitative way.

Table 1-1 summarizes the type of specifications which have been prepared for several recent or proposed systems (some of which may no longer be under active consideration).

In this table, three types of specifications are noted.

(1) System/subsystem MTBF/MTTR. An "X" in this column indicates that required limits on mean-time-between-failure and mean-time-to-restore are given at a system or subsystem level. (Example: MTBF of propulsion system.)

(2) Allowable MTBF/MTTR for various failure types. An "X" in this column indicates that specifications regarding the allowable frequency and duration of various classes of failures are given. (Example: MTBF of failures which block the guideway.)

(3) Other service requirements. A number in this column signifies that some other consideration of service is specified. Explanation of the type of measure is given in correspondingly numbered paragraphs subsequent to Table 1-1.

TABLE 1-1. EXAMPLES OF TYPES OF SPECIFICATIONS

System	System/ Subsystem MTBF/MTTR	Allowable MTBF/MTTR By Failure Type/Severity	Special Service Measure
Denver[128][a]	X	--	--
Bart[127]	X	--	--
Post Oak[126]	X	--	--
AIRTRANS[125]	X	X	--
Morgantown PRT[124]	?	--	1
Sea-Tac[123]	X	X	2
Standard Light Rail Vehicle[122]	X	X	--
MARTA[121]	X	X	--
Medium Capacity AGT[119]	X	--	3
Transit X-Way Revenue Line[120]	X	X	4

(a) Numbers refer to specific references in the bibliography.

(1) In this reference, other service requirements are stated for the maximum allowable system downtime, the allowable frequency of downtime events, and limits on maximum downtime per event. Similar specifications are given for degraded performance. This type of specification impacts the lower bound on system MTBF and the upper bound on MTTR.

(2) Availability calculations (classic form) are required and values are specified along with MTBFs and MTTRs.

(3) Additional constraints are incorporated to limit the number of delay incidents experienced by the average commuter in a year and to limit the allowable number of line blockage failures. This type of specification essentially establishes dependability requirements.

(4) This reference does not actually specify additional conditions, but does specify a failure categorization scheme (presumably for evaluation purposes) involving a weighted probability of occurrence and its impact on passenger minutes of delay. The probability weights range from 1 (impossible) to 6 (expected frequently). An "operational effectiveness" calculation is made by multiplying the frequency weight by the passenger minutes of delay expected to result. Hence, it is a measure of expected delay.

As can be seen from these examples, there are limited cases where some form of service requirement has been imposed with regard to desired service capabilities. In no case, however, is there any evidence that the values specified were derived from some analysis of actual, in-service performance perceptions by the passenger.

6. SYSTEM PERFORMANCE MEASURES USED BY SELECTED OPERATING SYSTEMS

This section presents the measures utilized by selected operating systems. In general, these examples illustrate the following points:

(1) System level performance measures do not make a distinction between failure-induced delays and delays from other causes.

(2) System level measures appear to be useful summaries of performance or productivity. They do not, however, provide the detail and insights needed to improve performance.

(3) Direct measurement of passenger delay is not done. Yet, all measures have a relationship to the passengers perception of service.

(4) In addition to system measures, detail failure data, maintenance data, and availability information on vehicles is maintained. Vehicle management represents the most significant variable in providing service and controlling costs.

(5) No measure used by these operating systems is adequate for specification purposes based on passenger perception of failure-induced delay effects.

6.1 Dallas-Ft. Worth Airport AIRTRANS System[71]

Service performance measurements for the AIRTRANS System are based upon a daily service report which logs categories of delay incidents. These incidents are weighted by a "service factor" which is a subjective rating of the severity of the delay incident. The exact formulation of these service factors is not known at this time. However, they appear to represent a delay magnitude or restorability of the system following a delay incident. There are also indications that passenger inconvenience is prominent in the service factor. For example, the worst delay incident is "station bypass", presumably because this requires the passenger to make another loop to reach his desired destination. Therefore, while quantitative measures of passenger delays or inconveniences are not taken, it appears that these measures are qualitatively handled in the AIRTRANS service measure.

At the end of each day, the weighted delay incident totals are combined into a daily service measure.

$$SM = \sum_i W_i N_i$$

where SM = Service measure

N_i = Number of delay incidents of the ith type

W_i = Weight associated with ith-type delay incident.

The delay incidents recorded are not necessarily failure related. They include also incidents such as passengers holding doors, etc. Data are available, however, which would permit the identification of "failure-induced" service measures.

A scale of goodness has been established for values of SM as follows:

SM	Quantitative Service
0-50	Excellent
50-100	Good
100-150	Satisfactory
150-200	Marginal
Over 200	Poor

While this measure of performance is nearly entirely qualitative, it does represent an easily assembled and informative management summary. In certain respects, it is better than a total system delay measure or classical availability measure because of its sensitivity to passenger inconvenience.

6.2 PATCO Lindenwold Line - Semi-Automatic Rapid Transit[70]

PATCO regularly evaluates on-time performance as a service measure. Delays up to four minutes in duration are not logged. Delays in excess of four minutes are logged, together with each annulled train and station bypass (necessary to maintain schedule). Ten missed stations are counted as one annulled train (since there are ten stations between end terminals). The total number of late or annulled trains is counted as trips not run on schedule, from which the percent trips on time can be calculated. This

measure is like a classical availability measure in that it is a measure of "something delivered"/"something promised". The units are trips rather than hours, values for the Lindenwold line run between 97 and 99 percent.

Like AIRTRANS, the PATCO service performance measure is not exclusively derived from equipment failure-induced delays but rather incorporates all delays, from all sources. Data are maintained, however, which would allow the distinction to be made.

In addition to the system level performance measure, detailed failure and availability data are maintained for individual vehicles.

6.3 Bay Area Rapid Transit (BART)[65]

BART has several measures of system performance, each of which has a particular purpose. Two of these relate to passenger perception of service.

(1) Number and rate of serious system delays (over 10 minutes)

(2) Number of trains arriving at the end of the line within a specified tolerance (like the PATCO measure).

Both measures do not separate the delays caused by equipment malfunction although data exist to permit such a separation.

Work is currently underway to permit a criticality judgment to be made regarding serious delays. This subjective factor would include such things as the delay duration of the affected trains and the impact of the delay on the remainder of the system. Many records are maintained regarding vehicle availability, reliability, and maintainability. Vehicle availability counts are made each morning (number of cars available/number of cars in active fleet).

BART is attempting to move toward tracking significant passenger delays by merging train departure delay data and electronic fare gate data to determine passenger delays.

6.4 Seattle-Tacoma International Airport (Sea-Tac)[132]

Sea-Tac maintains comprehensive failure, repair, and system restore time data. The system performance measure used is "system availability" defined as

$$SA = \frac{MTBF}{MTBF + MTTR}$$

where SA = system availability
MTBF = mean operating time between service interruption
MTTR = mean time to restore service to satellites
Failure = any incident which causes a delay in service to the satellite in excess of two minutes.

Sea-Tac has been running at approximately 0.996 with this measure compared to a goal of 0.998.

6.5 Morgantown PRT[133]

Detailed logs of performance are maintained for the Morgantown PRT system. Data on scheduled operating hours, actual operating hours, downtime events, passengers carried, and vehicle availability are gathered. The system performance measure utilized is termed "system dependability" defined as follows:

$$SD = (SA)(VA)(R)$$

where SA = System availability computed as the measured operating time/scheduled operating time ratio.

VA = Vehicle availability. This is <u>not</u> availability in the operational sense but rather a factor which represents the capacity capability of the system--presumably observed prior to starting service for the day.

$$VA = \frac{\text{Number of Vehicles Capable of Service}}{\text{Number Required for Service}}$$

VA has a maximum value of 1.

R = Trip reliability. The definition and computation of this term is unknown at the time.

This measure is of the classical dependability form and from a passenger viewpoint indicates the likelihood of being able to make a trip without incurring a delay.

Of the five examples discussed in this section, the Morgantown PRT is the only operating system monitoring a performance measure which directly relates to a classical conceptual measure of passenger perception.

APPENDIX

BIBLIOGRAPHY

This appendix contains a listing of the specific information sources used in this task. There is no significance to the order in which they are listed; the order listed reflects only the acquisition sequence. As an initial guide to those readers who wish to dig deeper into the subject of service availability, the specific references were categorized according to their relevance to the subject of service availability. This relevance is indicated by a "bullet" in the appropriate matrix column following the reference.

The meaning and organization logic behind the matrix column headings are discussed in the following expanded outline.

Service Availability citations contain information which directly treat some aspect of service availability as defined in this program. The particular emphasis of the reference is divided into

> Measures, indicating a reference dealing with the dimensions or appropriate expressions for service availability,
>
> Models, indicating a reference dealing with formulation and techniques for computing service availability, or
>
> Applications, indicating a reference dealing with actual use of some form of service availability expression. References in this category are further categorized according to the context in which such use was made:
>
> > Planning, indicating a use of service availability in the context of transit system planning activities,
> >
> > Specification, indicating a use of service availability considerations in the specification package for a transit system, or
> >
> > Measurement, indicating a use of service availability concepts in measuring the performance of operating transit systems or parts of transit systems.

Transit System Simulation Models citations contain information regarding modeling of transit system operations, particularly operations of PRT-type systems. In general, these models do not treat service availability directly. However, the models may prove to be useful in assessing the impacts of unavilability on transit system performance.

R/M Considerations citations contain information on reliability and maintainability aspects of systems and, hence, are indirectly related to the subject of service availability.

Passenger Perception citations deal with passenger surveys and analyses pertaining to transit service characteristics as perceived by the passenger and, hence, are important in establishing the nature of a service availability measure.

General citations contain information which is relevant in only a general way, not fitting into any of the above categories.

A reference which is not keyed to one of these categories does not contain information relevant to the study of service availability.

Reference	Service Availability: Measures	Service Availability: Models	Service Availability: Applications: Planning	Service Availability: Applications: Specification	Service Availability: Applications: Measurement	Transit System Simulation Models	R/M Considerations	Passenger Perception	General
(1) MODELS FOR ASSESSING TRIP DEPENDABILITY IN AUTOMATED GUIDEWAY TRANSIT NETWORKS, Kershner, D.L. and Roesler, W.J.(Johns Hopkins University/APL), Report No. CP 047/TPR 036, August 1976, 61pp.	●	●							
(2) OPERATIONAL AVAILABILITY AND RELIABILITY MODEL, Shulman, H.I. and Smith, H.L., IEEE Transactions on Reliability., Vol R-23, No.5, December 1974, pp 290-294.		●							
(3) A COST-BASED AVAILABILITY ALLOCATION ALGORITHM, McNichols, R.J. and Messer, G.H.,Jr.(Texas A&M University and U.S. Army Material Command Logistic Data Center), IEEE Transactions on Reliability, Vol R-20, No.3, August 1971, pp 178-182.		●							
(4) ON-LINE SYSTEM AVAILABILITY AND SERVICE SIMULATION(OLSASS), Doshay, I.(Acrojet ElectroSystems Company),IEE Transactions on Reliability, Vol R-20, No.3, August 1971, pp 142-147.		●							
(5) AVAILABILITY ANALYSIS OF DUAL MODE SYSTEMS, Toye, C.R.(Transportation Systems Center), Report No. UMTA-MA-06-0029-74-1, April 1974, for DOT/UMTA/Office of Research and Development,Contract UM408/R4764, PB 232 953, 23pp.	●	●							
(6) AVAILABILITY--CONCEPTS AND DEFINITIONS, Heimann, D. (Transportation Systems Center), Appendix B, 19pp.	●	●							
(7) PRELIMINARY STUDY OF RELIABILITY AND SIZING FOR AUTOMATED TRANSPORTATION COMPUTER SYSTEMS,Hunter, I.R.,Brown, N.K., McDowell, R.B.,Tolmei R., and Yingling, J.A.(Johns Hopkins University/APL), Report No. TIR 004, August 1974, for UMTA, Contract DOT-U-30010, 233pp.							●		
(8) RECOMMENDED METHODS OF ACHIEVING HIGHER METROLINER CAR AVAILABILITY, General Electric Company, Transit Systems Products Department,Final Report, May 1971, For DOT/FRA, Contract DOT-FR-10012, PB 202 036,157pp.					●				
(9) HOW TO RUN AN AUTOMATED TRANSPORTATION SYSTEM, Boyd, R.K. and Lukas, M.P.(TRW Systems Group, TRW Inc.), IEEE Transactions on Systems, Man, and Cybernetics, Vol SMC-2, No.3, July 1972, pp 331-341.						●			
(10) THE ACCOMMODATION OF ANOMALOUS EVENTS IN SYNCHRONOUS TRANSPORTATION NETWORKS, Boyd, R.K.,Kan, I.F. and Bogert, L.R.(TRW Systems Group, Washington Operations), 1971 IEEE International Convention Digest, March 22-25,1971,pp 420-421.									
(11) A COMPARATIVE REVIEW OF 23 URBAN TRANSPORTATION STUDIES: VOLUME II, Lieb,J.G. and Lenard, M.(The MITRE Corporation), Report No. MTR-6925, August 1975, For DOT/UMTA, Contract No. DOT-UT-50016, 191pp.									
(12) MORGANTOWN PROJECT VEHICLE SYSTEM RELIABILITY AND MAINTAINABILITY PROGRAM PLAN, Currence, C.F. and Ijanis, T.E.(The Boeing Company), Report No. D191-62101-1, Code Ident. No. 81205, 66pp.	●			●					

Reference	SERVICE AVAILABILITY: MEASURES	MODELS	APPLICATIONS: PLANNING	SPECIFICATION	MEASUREMENT	TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
(13) PASSENGER PSYCHOLOGICAL DYNAMICS, Solomon, K.M., Solomon, R.J. and Silien, J.S. Journal of Urban Transportation Corporation, for the ASCE and the Urban Transportation Research Council, DOT/UMTA, Report No.3, June 1968, 181 pp.								●	
(14) AIRPORT AND AIR SERVICE ACCESS, de Neufville, R., Wilson, N. More. H.III, Gelerman, W., Landau, U. and Yaney, J.(Massachusetts Institute of Technology) Report No. DOT-TSC-OST-73-3, March 1973, For DOT/OST/Office of R&D Policy, PB 220 646, 146pp.									
(15) STUDY IN NEW SYSTEMS OF URBAN TRANSPORTATION--A STUDY OF COMMAND AND CONTROL SYSTEMS FOR URBAN TRANSPORTATION, General Electric Company, Transportation Systems Division, February 1968, for HUD, PB 178 281, 357pp.									
(16) MARKETING TECHNIQUES AND THE MASS TRANSIT SYSTEM, National Urban League and Mark Battle Associates, Report No. UMTA-DC-06-0020-73-1, July 1973, For UMTA/DOT, Contract DOT-UT 533, PB 223 735, 97pp.									
(17) PERFORMANCE AND GENERAL SYSTEM SPECIFICATIONS FOR THE REGIONAL FIXED GUIDEWAY SYSTEM, Simpson and Curtin and Midwest Planning and Research, Inc. January 1974, For The Twin Cities Area Metropolitan Transit Commission, Project No. MN-09-0008, 73pp.			●	●					
(18) SELECTED PAPERS FROM THE CONFERENCE ON DUAL-MODE TRANSPORTATION, Washington, D.C., May 29-31, 1974.									
(19) PERSONAL AND GROUP RAPID TRANSIT SYSTEMS AND TECHNOLOGIES: A SELECTED BIBLIOGRAPHY, Henderson, C. and Erickson, L.(Stanford Research Institute), May 1973, SRI Project PYU-2201, For NASA, Contract NASw-2455, 41pp.									
(20) SURVEY OF PRT VEHICLE MANAGEMENT ALGORITHMS, Priver, A.S.(Transportation Systems Center), Report No. UMTA-MA-06-0031-74-13, September 1974, For DOT/UMTA/Office of Research And Development, PB 238 942, 100pp.						●			
(21) NETWORK MODEL STUDIES FOR PERSONAL RAPID TRANSIT, Kershner, D.L. and Roesler, W.J.(Johns Hopkins University/APL), Report No. CP 040A/TPR 031A, June 1975, For DOT/UMTA, Contract DOT-UT-30010, 76pp.						●			
(22) NETWORK MODEL STUDIES FOR AUTOMATED GUIDEWAY TRANSIT: ADVANCED GROUP RAPID TRANSIT MODELS, Kershner, D.L., Rand, R.C. and Roesler, W.J.(Johns Hopkins University/APL), Report No. CP 044/TPR 033, February 1976, For DOT/UMTA, Contract DOT-UT-30010, 187pp.						●			
(23) SYSTEMS ANALYSIS OF URBAN TRANSPORTATION, VOLUME I: SUMMARY, General Research Corporation, January 1968, PB 178 261, For HUD, 223 pp.									
(24) STUDYING TRANSPORTATION SYSTEMS FROM THE CONSUMER VIEWPOINT: SOME RECOMMENDATIONS, Hille, S.J., Brunner, G.A., Paine, F.T. and Nash, A.N.(University of Maryland), September 1967, For Department of Commerce/Bureau of Public Roads, 13pp.								●	

Reference	SERVICE AVAILABILITY: MEASURES	SERVICE AVAILABILITY: MODELS	SERVICE AVAILABILITY: APPLICATIONS: PLANNING	SERVICE AVAILABILITY: APPLICATIONS: SPECIFICATION	SERVICE AVAILABILITY: APPLICATIONS: MEASUREMENT	TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
(25) THE LINDENWOLD LINE: MODEL FOR EFFICIENT RAIL RAPID-TRANSIT, Dallaire, E.E.(Editor, Civil Engineering), Civil Engineering-ASCE, September 1970, 5pp.									●
(26) SUMMARY DATA FOR SELECTED NEW URBAN TRANSPORTATION SYSTEMS, Casey, R.F. (Transportation Systems Center), Report No. DOT-TSC-OST-72-35, November 1972, For OST/Office of Systems Analysis and Information, 152 pp, PB 219 254.									
(27) DEVELOPMENT OF PERFORMANCE SPECIFICATIONS FOR THE REGIONAL FIXED GUIDEWAY SYSTEM, Phase III-A-2, Simpson and Curtin, and Midwest Planning and Research, Inc., January 1974, For The Twin Cities Area Metropolitan Transit Commission, 103pp.									
(28) ANALYSIS AND SIMULATION OF ALTERNATE TRANSIT SYSTEMS FOR THE TWIN CITIES, Honeywell Inc/Urban and Environmental Systems, Reoort No. 74-4, January 1974, Project No. MN-09-0008, For the Twin Cities Area Metropolitan Transit Commission, 79pp.									
(29) DEVELOPMENT OF PERFORMANCE SPECIFICATIONS FOR THE REGIONAL FIXED GUIDEWAY SYSTEM: SUMMARY REPORT, Simoson and Curtin, and Midwest Planning and Research, Inc., January 1974, 30 pp.									
(30) ANALYSIS AND SIMULATION OF ALTERNATE TRANSIT SYSTEMS FOR THE TWIN CITIES: TECHNICAL SUPPLEMENT, Honeywell Inc./Urban and Environmental Systems, Report No. 74-4-A, January 1974, For the Twin Cities Area Metrooolitan Transit Commission, 169pp.									
(31) THE EFFECTS OF TRANSPORT IMPROVEMENTS ON URBAN DEVELOPMENT: TESTS ON AN OPPORTUNITIES MODEL, Brindle, R.E.(University of Melbourne), Special Report No.6, September 1970, 181pp.									
(32) SYSTEMS RELIABILITY IN QUALITY CONTROL, DePauw, J.R.(Arthur Anderson & Company), Quality Control in Rail Transoortation, Railway Systems and Management Association, August 1964, pp 23-26.									●
(33) RAIL TRIP TIME RELIABILITY: EVALUATION OF PERFORMANCE MEASURES AND ANALYSIS OF TRIP TIME DATA, Martland, C.D.(Massachusetts Institute of Technology), Volums 2 Thru 12, June 1972, PB 244 122, PB -244 125, PB 244 127.	●								
(34) A MATHEMATICAL PROGRAMMING MODEL FOR OPTIMAL SCHEDULING OF BUSES' DEPARTURES UNDER DETERMINISTIC CONDITIONS, Friedman, M.(Arizona State University), Transportation Research, Vol 10, 1975, pp 83-90.									
(35) MEASUREMENT AND EVALUATION OF TRANSPORTATION SYSTEM EFFECTIVENESS, Pardee, F.S. et al,(The RAND Corporation), Memorandum RM-5869-DOT, September 1969, Supported by FRA/OHSGT, Contract DOT-FR-9-0054, 463pp.								●	
(36) RELIABILITY IN RAIL RAPID TRANSIT AND COMMUTER CARS, Line, R.H., and Cord, J.M.(The Boeing Company/Vertol Division), 7op.					●				
(37) COST COMPARISON OF BUSWAY AND RAILWAY RAPID TRANSIT, Miller, D.R., Goodwin, B.C., Hoffman, G.A. and Holden, W.H.T.(Daniel, Mann, Johnson, and Mendenhall, Inc.) Highway Research Record No.459, "Evaluation of Bus Transit Strategies," 52nd Annual Meeting of the TRB, 1973, pp 1-12.									●

Reference	SERVICE: MEASURES	SERVICE: MODELS	AVAILABILITY APPLICATIONS: PLANNING	AVAILABILITY APPLICATIONS: SPECIFICATION	AVAILABILITY APPLICATIONS: MEASUREMENT	TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
(38) LINDENWOLD RAIL LINE AND SHIRLEY BUSWAY: A COMPARISON, Vuchic, V.R., and Stanger, R.M.(University of Pennsylvania),Ibid (37), pp 13-28.									●
(39) EVALUATION OF INTERSECTION-DELAY MEASUREMENT TECHNIQUES,Sofokdis,H., Tilles, D.L. and Geiger, D.R.(FRA), Highway Research Record No. 453, "Traffic, Capacity, Delay, and Quality of Service", 1973, 53rd and 54th Annual Meetings of the GRB, pp 28-55.									
(40) FAULT TREE ANALYSIS FOR SYSTEMS RELIABILITY, Crosetto, P.A.(Douglas United Nuclear, Inc), 14th ISA Power Instrumentation Symposium, N.Y.,N.Y., 1971, 5pp.									
(41) RELIABILITY ENGINEERING: IMPACT ON TRANSPORTATION CONTROL SYSTEMS, Schwartz, H.C.(General Railway Signal Company), Railway Systems Controls, July/August 1974, pp 15-24.				●			●		
(42) CONDENSATION OF THE REPORT ON THE TESTING & EVALUATION OF THE TRANSIT EXPRESSWAY, MPC Corporation, Project No. PA-MTD-2, June 20,1967, For the Port Authority of Allegheny County and HUD. 50pp.									
(43) REPORT ON TESTING AND EVALUATION OF THE TRANSIT EXPRESSWAY, MPC Corporation, Mass Transportation Demonstration Project No. PA-MTD-2, February 20,1967, For The Port Authority of Allegheny County, and HUD, Contract 602, 271pp.									
(44) MODAL-CHOICE MODEL FOR RELATING DEMAND TO INVESTMENT, Bellomo, S.J., Turner, G., and Johnston, D.K.(Alan M. Voorhees and Associates,Inc), Highway Research Record, No 392, "Transportation Demand and Analysis Techniques", 51st Annual Meeting of the HRB, 1972, pp1-1-25.									
(45) TRANSPORTATION POLICIES AND THE DELIVERY OF SOCIAL SERVICES IN A SMALL CITY, Kidder, A.E.(North Carolina Agricultural and Technical State University), Transportation Research Record, No.516,"Transportation for the Poor, the Elderly, and the Disadvantaged", 53rd Annual Meeting of the TRB, 1974, pp 21-27.									
(46) A PROPOSED STANDARD FOR TRANSPORTATION SYSTEMS, Aronstein, R.H.(Computer Sciences Corporation), Proceedings of the Fourth Annual Intersociety Conference on Transportation, July 18-23, 1976, pp 1-9.								●	
(47) THE MORGANTOWN PRT ALTERNATIVES ANALYSIS STUDY: AN APPLICATION OF AN EVALUATION TECHNIQUE, Miller,D.R.(Barton-Aschman Associates,Inc.), 1975 International Conference on Personal Rapid Transit, Volume I, September 16-19,1975,Regency Hotel, Denver Colorado, Universities of Minnesota and Colorado, 15pp.									
(48) THE DEVELOPMENT OF A MODEL FOR ANALYSIS OF THE COST EFFECTIVENESS OF ALTERNATIVE TRANSIT SYSTEMS, Anderson, J.E.(Univ. Minnesota), Raytheon Company, 42pp, Ibid (47).									
(49) A SCENARIO-SPECIFIC METHODOLOGY FOR THE EVALUATION OF PERSONAL RAPID TRANSIT ALTERNATIVES AND NETWORK DESIGNS, Iverson, S.C.(California Polytechnic State University), 19pp, Ibid (47).									

Reference	SERVICE AVAILABILITY: MEASURES	SERVICE AVAILABILITY: MODELS	APPLICATIONS: PLANNING	APPLICATIONS: SPECIFICATION	APPLICATIONS: MEASUREMENT	TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
(50) A RATIONALE FOR AUTOMATED PERSONAL AND GROUP TRANSIT DEVELOPMENT, Lay, R.K.(The MITRE Corporation), 18pp, Ibid (47).									
(51) THE INFLUENCE OF THE SIZE AND NUMBER OF VEHICLES ON THE PERFORMANCE AND SERVICE QUALITY OF GROUP RAPID TRANSIT SYSTEMS, Bahm, G.(Institut Fur Verkehrswesen, Univ. of Karlsruhe), 21 pp, Ibid (47).						●			
(52) AN EVALUATION OF GROUP RAPID TRANSIT, Ford, B.M., Waddell, M.C. and Williams, M.B.(The Johns Hopkins University/APL), 17pp, Ibid (47).						●			
(53) DEVELOPMENT AND EVALUATION OF SERVICE POLICIES FOR MEDIUM HEADWAY AUTOMATED RAPID TRANSIT SYSTEMS, Thangavelu, K.(De Leuw, Cather and Company), 26pp, Ibid (47).						●			
(54) OPERATIONAL ISSUES IN SMALL VEHICLE AUTOMATED GUIDEWAY SYSTEMS, Anderson, P.A. and Pejsa, J.H.(Urban and Environmental Systems, Honeywell, Inc.), 18 pp, Ibid (47).									
(55) VEHICLE MANAGEMENT ON LARGE PRT NETWORKS, Irving, J.H., Bernstein, H., Katz, J., Dergarabedian, P. and Silva, T.H.(The Aerospace Corporation), 44pp, Ibid (47)						●			
(56) SYSTEMS MANAGEMENT ANALYSIS OF LARGE AGT NETWORKS, Ross, M.S. and Melgaard, A.D.(IBM, Federal Systems Division and Advanced Transportation Systems), 15 pp, Ibid (47) Volume II.						●	●		
(57) CVS(COMPUTER-CONTROLLED VEHICLE SYSTEM), Ishi, T., Iguchi, M., and Koshi, M.(University of Tokyo), 10pp, Ibid (47) Volume II.									
(58) SOME RELIABILITY, DEPENDABILITY AND SAFETY CONSIDERATIONS FOR HIGH-CAPACITY PRT SYSTEMS, Olson. C.L., Fuller, G.H. and Fling, R.B. (The Aerospace Corporation), 19pp, Ibid (47) Volume II.									●
(59) SERVICE DEPENDABILITY EVALUATION AND DESIGN CONSIDERATIONS FOR AUTOMATED TRANSIT SYSTEMS, Ems, A.(TRW/Systems Management Contractor), 18 pp, Ibid (47), Volume II.									
(60) NETWORK DYNAMICS IN A LARGE PERSONAL RAPID TRANSIT SYSTEM FAILURE, Wade, R.M.(International Business Machines Corp.), 13pp, Ibid (47), Volume II.						●			
(61) THE ECONOMICS OF HIGH-CAPACITY PRT SYSTEMS, Bush, L.R.(The Aerospace Corporation), "Personal Rapid Transit", A collection of Papers on a New Type of Urban Transportation, April 1972, pp 231-260.									
(62) EMERGENCY STRATEGIES FOR SAFE CLOSE-HEADWAY OPERATION OF PRT VEHICLES, Bernstein, H. and Schmitt, A.(The Aerospace Corporation), pp 351-360, Ibid (61).									
(63) VEHICLE MANAGEMENT FOR PRT SYSTEMS, Ford, B.M., Roesler, W.J. and Waddell, M.C.(Johns Hopkins University/APL), National Conference on Personal Rapid Transit, November 1-3, 1971, University of Minnesota, 22pp.						●			
(64) COMPUTER SIMULATION TECHNIQUES, Naylor, T.H., Balintfy, J.L., Burdick, D.S. and Chu, K.(Duke University and Tulane University), John Wiley & Sons, Inc., (C)1966, 352pp.									

Reference	Service Availability: Measures	Service Availability: Models	Service Availability: Applications: Planning	Service Availability: Applications: Specification	Service Availability: Applications: Measurement	Transit System Simulation Models	R/M Considerations	Passenger Perception	General
(65) BART RELIABILITY & AVAILABILITY DATA EXPERIENCE, King, J.H.(Mgr. of Reliability Engineering-BART), Presentation at the 1976 UMTA AGT Service Availability Workshop, 17pp.					●				
(66) NEW SYSTEMS REQUIREMENT ANALYSIS PROGRAM, UMTA TRANSPORTATION PLANNING SYSTEM(UTPS), SKETCH PLANNING WORK-SHOP PROCEEDINGS, Reoort No.UMTA-RDD-72-5, October 19,1972, PB 214 263, 52pp.									
(67) TOWARD THE DEVELOPMENT OF MEASURES OF CONVENIENCE FOR TRAVEL MODES, Stopher, P.R., Spear, B.D. and Sucher, P.O.(Cornell University), For HRB Annual Meeting of Jan.1974, August 1973., 41pp.								●	
(68) RELIABILITY OF EQUIPMENT,Vollmar, J.R.(Louis T. Klauder and Associates), Presented at the APTA Conference:Car Equioment and Ooerations Session, Aoril 3,1975, 5pp.									●
(69) ANALYSIS OF DUAL MODE SYSTEMS IN AN URBAN AREA; VOLUME I: SUMMARY, Benjamin, P. Et al, Reoort No. DOT-TSC-OST-73-16A,I, December 1973, DOT/TSC/OST/Office of Systems Engineering, 40pp.									
(70) AN OPERATOR'S EXPERIENCE WITH SEMI-AUTOMATED RAPID TRANSIT, Vigrass, J.W.(Port Authority Transit Corp. of Pa. and N.J.), Presented at the UMTA Worksop on Automated Guideway Transit Service Reliability, Oct. 6,1976,Andover, Mass.,13po+Tabs.					●				
(71) OPERATOR EXPERIENCE DESCRIPTION OF THE AIRTRANS SYSTEM, Morgan, J.D., Ibid (70), 8pp and Tabs.					●				
(72) SAFETY AND RELIABILITY OF AUTOMATED URBAN TRANSPORTATION SYSTEMS, Heinrich, K.,(Industrienlagen-Betriesbsgesellschaft mbH), Ibid (70), 12pp.									●
(73) THE CHRONICLE OF RELIABILITY, Grose, V.L.(Tustin Institute of Technology), Preoared under the Auspices of the AIAA System Effectiveness Technical Committee, June 3,1968, 115pp.							●		
(74) AN ECONOMIC PERSPECTIVE ON PRT NETWORKS, Godfrey(Ohio State University), MB, Wilson, N.H.M., and Yamamoto, W.M.(MIT), 1973 International Conference on PRT, Progress, Problems and Potential, Personal Raoid Transit II, February 1974, pp 211-230.									
(75) ECONOMIC AND PERFORMANCE EVALUATIONS OF OPERATING POLICIES FOR A CLOSED-LOOP PRT NETWORK, Godfrey, M.B.(Ohio State University), Ibid (75),pp 479-498.									
(76) AN ANALYSIS OF MINIMUM SAFE HEADWAY FOR NO COLLISIONS, Lobsinger, D.J., (Transportation Technology, Inc.-Otis Elevator Company), Ibid (75), pp 391-398.									
(77) REPORT ON TESTING AND EVALUATION OF THE TRANSIT EXPRESSWAY, MPC Corporation, Port Authority of Allegheny County and HUD, Project No. PA-MTD-2, Contract 602, February 20,1967, 272pp.									
(78) REPORT ON TRANSIT EXPRESSWAY PHASE II, MPC Corooration, For Port Authority of Allegheny County and DOT, Project No. PA-06-0009, Contract H-602, June 1973, 266pp.									

Reference	Service Availability: Measures	Service Availability: Models	Applications: Planning	Applications: Specification	Applications: Measurement	Transit System Simulation Models	R/M Considerations	Passenger Perception	General
(79) ADVANCED GROUP RAPID TRANSIT SYSTEM DEVELOPMENT PROGRAM(Executive Summary Report), Boeing Vertol Company, September 1975, For DOT/UMTA, Contract DOT-UT-50012, Report No. D336-10029-1, Project No. PA 060032, 124pp.									
(80) HIGH PERFORMANCE PERSONAL RAPID TRANSIT SYSTEM: RELIABILITY PLAN, Otis Elevator Company/Transportation Technology Div, Report No. HPPRT OTIS-TTD-080, July 1975;Interim Report, For DOT/UMTA, Contract No. DOT-UT-50011, 68pp.			●						
(81) SAFETY ANALYSIS AND FAILURE MODE STUDY REPORT, Jines, R.H.(Boeing Vertol Company), Report No. D336-10026-1, July 31,1975, For DOT/UMTA, Contract DOT-UT-50012, 49pp.									
(82) HIGH PERFORMANCE PERSONAL RAPID TRANSIT(HPPRT):INTERFACE CONTROL DOCUMENT, Ross, J.L.(Boeing Vertol Company), Report No. D336-10007-1, April 16,1975, For DOT/UMTA, Contract DOT-UT-50012, Project No. PA 060032,28pp.									
(83) ADVANCED GROUP RAPID TRANSIT SYSTEM(AGRTS): URBAN DEPLOYABILITY STUDIES, Stuverude, H.N.(Boeing Vertol Company), Report No. D336-10022-2, August 1975, For DOT/UMTA, 147pp.			●						
(84) SOFTWARE REQUIREMENTS SPECIFICATION FOR THE HIGH-PERFORMANCE PERSONAL RAPID TRANSIT SYSTEM(HPPRT),Lawrence,G.M.(Boeing Vertol Company), Doc. No. S336-10201-1, 49pp.									
(85) NOTES ON VALIDATING/VERIFYING COMPUTER SIMULATION MODELS, Berman, M.B., The Rand Corporation, No. P-4891, August 1972, Presented at the SHARE Conference , SHARE System Simulation Project, Toronto, Canada, 11pp.									
(86) OPERATING POLICIES FOR PERSONAL RAPID TRANSIT, Morag,D.(DOT), Report No. UMTA-RDD-8-74-2, May 1974, For UMTA/DOT, PB 239 825, 92pp.									
(87) PERSONAL RAPID TRANSIT RESEARCH CONDUCTED AT THE AEROSPACE CORPORATION, Olson, C.L. and Berstein(The Aerospace Corporation), Report No. UMTA-CA-06-0071-76-1, 18 March 1976, for UMTA/DOT, Contract No. DOT-UT-60006, PB 256 846, 276pp.	●		●						
(88) PUBLIC TRANSPORTATION RESEARCH NEEDS, Highway Research Board Special Report No. 137, National Research Council, National Academy of Sciences, 55pp.									
(89) DEMAND-RESPONSIVE TRANSPORTATION SYSTEMS AND SERVICES, Special Report No. 165, Transportation Research Board, National Research Council,177pp.									
(90) USE OF ECONOMIC, SOCIAL, AND ENVIRONMENTAL INDICATORS IN TRANSPORTATION PLANNING, Highway Research Record, No. 410, HRB/NRC/NAS/NAE, 51st Annual Meeting of the HRB, 1972, 70 pp.									
(91) NEW SYSTEMS REQUIREMENTS ANALYSIS PROGRAM TRANSPORTATION SYSTEM EVALUATION INDICATORS, Peat, Marwick, Mitchell and Co., Report No. UMTA-IT-06-0050-73-1, May 1973, For UMTA/DOT,Contract IT-06-0050, PB 221 572, 98pp.									

Reference	SERVICE AVAILABILITY: MEASURES	SERVICE AVAILABILITY: MODELS	SERVICE AVAILABILITY APPLICATIONS: PLANNING	SERVICE AVAILABILITY APPLICATIONS: SPECIFICATION	SERVICE AVAILABILITY APPLICATIONS: MEASUREMENT	TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
(92) AUTOMATED GUIDEWAY NETWORK TRAFFIC MODELING, Toye, C.R.(Transportation Systems Center), Reoort No. DOT-TSC-OST-72-7, February 1972, PB 211 956, 25pp.						●			
(93) HEADWAY SEPARATION ASSURANCE SUBSYSTEM(HSAS), Evans, R.R. and Cowes, K.(Alden Self-Transit Systems Corporation), Reoort No. UMTA-MA-06-0031-75-4, July 1975, for DOT/UMTA, Contract DOT-TSC-421,PB 244 667, 118pp.									
(94) MACROANALYSIS OF THE IMPLICATIONS OF MAJOR MODAL SHIFTS IN INTEGRATED REGIONAL TRANSPORTATION NETWORKS, Billheimer, J.W.,Bullemer, R. and Holoszyc,M.(SYSTAN,Inc), Report No. DOT-OS-50265-S, April 1976, For DOT/Office of R&D Policy, Contract DOT-OS-50265, PB 256 137, 113pp.									
(95) NEW SYSTEMS REQUIREMENTS ANALYSIS PROGRAM: A PROCEDURE FOR LONG RANGE TRANSPORTATION(SKETCH) PLANNING--TECHNICAL REPORT, DOT/UMTA, Report No. 3-RDD-7, July 1973, PB 223 344, 31pp.									
(96) EVALUATION OF THE WABCO AC PROPULSION SYSTEM: VOLUME 2, RELATED REPORTS AND EXHIBITS, Hoppe, C.F. Et Al(Cleveland Transit System), Report No. UMTA-OH-06-0006-74-4, September 1974, For DOT/UMTA, Contract No. OH-06-0006 , PB 245 390, 145pp.									
(97) WABCO DATA ACQUISITION SYSTEM: AC PROPULSION PROJECT, Smith,R.D. et al.(Cleveland Transit System), Report No. UMTA-OH-06-0006-73-1, August 1973, For UMTA/DOT, Contract OH-06-0006, PB 223 898, 102pp.									
(98) EVALUATION OF THE WABCO AC PROPULSION SYSTEM: VOLUME I; CTS/UMTA AC PROPULSION PROJECT, Hoppe,C.F. et al,(Cleveland Transit Systems), Report No. UMTA-OH-06-0006-74-3, Seotember 1974, For DOT/UMTA, Contract OH-06-0006, PB 245 389, 102pp.									
(99) NEW SYSTEMS REQUIREMENTS ANALYSIS PROGRAM, PRC Systems Services Co., Report No. PRC/SSC 1.1, July 7,1972, For DOT/UMTA, Contract DOT-UT-20021, PB 218-862, 130pp.									
(100) A SIMULATION MODELING SYSTEM FOR THE STUDY OF SCHEDULING, ROUTING, AND SERVICE CONTROL IN TRANSPORTATION NETWORKS, Davis, E.A.(APL/JHU), Presented at the 38th National Meeting of the ORSA, October 28,1970. 25pp.		●							
(101) SIMULATION AND ANALYSIS OF A DEMAN RESPONSIVE, AUTOMATED TRANSIT SYSTEM, McGean, T.J. and Chambliss, A.G.(DeLeuw Cather & Company, and The MITRE Corporation), Journal of Aircraft, Vol 12, No.10, October 1975, pp 769-772.									
(102) A DEMAND-RESPONSIVE TRANSPORTATION SIMULATION PROGRAM: PROGRAM SUMMARY AND USER GUIDE, Chambliss, A.G.(The MITRE Corporation), WP-10726, September 9,1974, For DOT/UMTA, Contract DOT-UT-10028, 37pp.						●			
(103) A FIXED GUIDEWAY TRANSPORTATION SIMULATION PROGRAM, Chambliss, A.G. The MITRE Corporation, Report No. MTR-6919, May 1975, For UMTA/DOT, Contract DOT-UT-50016, 57pp.						●			

Reference	SERVICE AVAILABILITY					TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
			APPLICATIONS						
	MEASURES	MODELS	PLANNING	SPECIFICATION	MEASUREMENT				
(104) SYSTEMS ANALYSIS OF PERSONAL RAPID TRANSIT SYSTEMS, Thangavelu, K., Doctoral Thesis, Northwestern University, 1974, 234pp.									
(105) DUAL MODE TRANSIT SYSTEM, PHASE I, Transportation Technology, Inc., Report No. UMTA-CO-06-0006-74-1, June 1974, For UMTA/DOT, Contract No. CO-06-0006, PB 239 841, 312 pp.			●						
(106) CHARACTERISTICS OF URBAN TRANSPORTATION SYSTEMS: A HANDBOOK FOR TRANSPORTATION PLANNERS, DeLeuw, Cather & Company and the Urban Institute, Report No. UMTA-IT-06-0049-75-1, May 1975, For UMTA/DOT, Contract No. IT-06-0049, PB 245 809/233-580/AS, 170pp.									
(107) DUAL MODE TRANSIT SYSTEM, GM Transportation Systems Div., General Motors Corporation, GM Technical Center, Report No. UMTA-NI-06-0016-74-1, July 1974, For DOT/UMTA, PB 238 941, 462pp.									
(108) SAFETY AND RELIABILITY OF AUTOMATED SYSTEMS IN LOCAL GUIDED PUBLIC TRANSPORT, Heinrich, K.(Industrieanlagen Betriebsgesellschaft mbH), Public Transport and People/International Conference, Paris, April 16-18,1975, pp 175-177									
(109) TRANSIT OPERATING PERFORMANCE AND THE IMPACT OF THE SECTION 5 PROGRAM, U.S. Department of Transportation, UMTA/Office of Policy and Program Development, November 1976, 81pp.									
(110) DUAL MODE TRANSIT SYSTEM AVAILABILITY STUDY, Eugene L, Bagenstos,L., and Farley, J.,(Transportation Technology,Inc.) May 1974, For DOT/UMTA, Contract No. DOT-UT-40009, 112 pp.			●						
(111) CHESAPEAKE MASS TRANSPORTATION DEMONSTRATION PROJECT-VA-MTD-1, Final Report, Wilbur Smith and Associates, January 1969, For the City of Chesapeake, Va., PB 185 657, 73pp.									
(112) STATIONS, SYSTEM PERFORMANCE AND SCHEDULING IN PERSONAL RAPID TRANSIT NETWORKS, Sirbu, M.A.,Jr., Center for Transportation Studies/MIT, CTS Report No. 74-2, March 1974, 260pp.									
(113) MODELS, ANALYSIS, AND APPROXIMATIONS FOR SYSTEM RELIABILITY AND AVAILABILITY, Messinger, M.(Polytechnic Institute of Brooklyn) 1967, Dissertation No. 67-13,408, 291pp.		●							●
(114) A SIMULATION AND ANALYSIS OF A LARGE VEHICLE DEMAND RESPONSIVE PRT SYSTEM, FEATURING OFFLINE STATIONS WITH VEHICLE STORAGE CAPACITIES, (Microfilm), Chambliss, A.G., Master's Thesis, Howard University, Wash. D.C. 1974.						●			
(115) TRANSPORTATION PROGRAMMING PROCESS, Proceedings of the Conference March 23-26, 1975 of the Transportation Research Board, Special Report No. 157, (see "Evaluation", pp67-72),									

Reference	SERVICE AVAILABILITY: MEASURES	SERVICE AVAILABILITY: MODELS	APPLICATIONS: PLANNING	APPLICATIONS: SPECIFICATION	APPLICATIONS: MEASUREMENT	TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
(116) CHOICE OF TRANSPORT TECHNOLOGY UNDER VARYING FACTOR ENDOWMENTS IN LESS DEVELOPED COUNTRIES, Roberts, P.(Charles River Associates, Inc), Report No. CRA-3-138-30, June 1970, For Office of the Assistant Secretary for Policy and International Affairs/DOT, PB 195 291, 354pp.									
(117) RELIABILITY STUDIES AT THE SNCF, Autruffe, H., Review Generale de Chemins de Fer, Vol 94, January 1975, 20pp,(IN FRENCH).									
(118) AN APPROACH TO THE METROLINER UPDATING PROGRAM, Uher, R.A.(Westinghouse Electric Corporation), Conference Paper No. C-75-354-6-1A, March 1975, IEEE Land Transportation Committee of the IEEE Industry Application Society, 1975 Joint ASME/IEEE Railroad Conference. April 13-16,1975. 18pp.									
(119) GENERAL SYSTEM REQUIREMENTS FOR A MEDIUM CAPACITY AUTOMATED GUIDEWAY TRANSIT SYSTEM, Prepared by the User Requirement Subcommittee, Automated Guideway Transit Task Force, American Public Transit Association, December 1975, pp 13-17.				●					
(120) TRANSIT EXPRESSWAY REVENUE LINE: CONTRACT BOOK; FOR PROCUREMENT AND INSTALLATION OF A VEHICLE AND TRAIN CONTROL AND COMMUNICATIONS, Kaiser Engineers, 1974, For the Port Authority of Allegheny County, Contract No. 1Z 5062, 18pp.				●					
(121) TECHNICAL PROVISION FOR PROCUREMENT OF TRANSIT VEHICLES, Part II, Parsons Brinckerhoff-Tudor-Bechtel, July 28,1975, For The Metropolitan Atlanta Rapid Transit Authority, Section 15.0 System Support, 23pp.				●					
(122) STANDARD LIGHT RAIL VEHICLE SPECIFICATION: TECHNICAL SECTION, Prepared by The Massachusetts Bay Transportation Authority and the San Francisco Municipal Railway, October 1972, Sponsored by DOT/UMTA, PB 220 748.				●					
(123) SPECIFICATIONS FOR SEATTLE-TACOMA INTERNATIONAL AIRPORT SATELLITE TRANSIT, The Richardson Associates , Beverly A. Travis Associates, And Battelle Columbus Laboratories, Jan 31,1969, For the Port of Seattle Commission.				●					
(124) MORGANTOWN PROJECT REQUIREMENTS AND CONSTRAINTS, Jet Propulsion Laboratory, Haglund, H.H.(Project Manager), No. 1200-17, December 21,1970.				●					
(125) PERFORMANCE SPECIFICATIONS FOR AIRTRANS(AIRPORT TRANSPORTATION SYSTEM), AT THE DALLAS-FORT WORTH REGIONAL AIRPORT, Dallas-Fort Worth Regional Airport Board, Battelle Columbus Laboratories, and Operations Research, March 15,1971.				●					
(126) POST OAK TRANSIT SYSTEM: Contract Documents and Specifications, Lockwood, Andrews and Newman, Inc. For the City of Houston, Texas, Special Consultant-Battelle Columbus Laboratories.				●					
(127) BART STANDARD SPECIFICATIONS: USER AND PROGRAMMER MANUALS, CONTRACT BOOKS, VEHICLE PROCUREMENTS, AUTOMATIC TRAIN CONTROL AND PROPULSION, AND CONSTRUCTION DETAILS, Parsons, Brinckerhoff: Tudor:Bechtel, July 1967, For the San Francisco Bay Area Rapid Transit Company, Contract No. 2Z4601.				●					
(128) URBAN DEMONSTRATION OF PERSONAL RAPID TRANSIT SYSTEM, By Urban Mass Transportation Administration, November 1972, 100pp.									

Reference	SERVICE AVAILABILITY: MEASURES	SERVICE AVAILABILITY: MODELS	SERVICE AVAILABILITY: APPLICATIONS: PLANNING	SERVICE AVAILABILITY: APPLICATIONS: SPECIFICATION	SERVICE AVAILABILITY: APPLICATIONS: MEASUREMENT	TRANSIT SYSTEM SIMULATION MODELS	R/M CONSIDERATIONS	PASSENGER PERCEPTION	GENERAL
(129) AUTOMATED GUIDEWAY TRANSIT TECHNOLOGY PROGRAM REVIEW, For the Automated Guideway Transit Technology Program, September 1976, For the U.S. Department of Transportation, Urban Mass Transportation Administration, Office of Research and Development.									
(130) SAFETY AND SYSTEM ASSURANCE REVIEW OF THE METROPOLITAN ATLANTA RAPID TRANSIT AUTHORITY(MARTA): RAIL RAPID TRANSIT SYSTEM DEVELOPMENT PROGRAM, The Office of Transit Management, Urban Mass Transportation Administration, Department of Transportation, April 1975, For The Metropolitan Atlanta Rapid Transit Authority.				●					
(131) Contract Documents: FURNISHING AND INSTALLING CAB SIGNAL AND INTERLOCKING EQUIPMENT FOR THE NORTH-SOUTH RAPID TRANSIT ROUTE, Spec. No. CYA-3737-72, Chicago Transit Authority, Chicago, Ill., 1974,97pp.									
(132) SEATTLE-TACOMA INTERNATIONAL AIRPORT: SEATTLE TRANSIT SYSTEM, Bitto, M.K.(Seattle/Tacoma Airport), Presentation at the 1976 UMTA AGT Service Availability Workshop, October 6,1976, Andover, Mass.,9pp + Ta					●				
(133) MPRT O&M PHASE OPERATING, AVAILABILITY AND MAINTENANCE HISTORY, Stone, A.L.(The Boeing Company), Document No. D191-62000-1, January 4,1977, 91pp+tabs.					●				

PART 2

THE DEVELOPMENT OF MEASURES OF SERVICE AVAILABILITY

TASK 2. FIELD INTERVIEWS

Contract No. DOT-TSC-1283

to

DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER

by

R. D. Leis

BATTELLE
Columbus Laboratories
505 King Avenue
Columbus OH 43201

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	2-1
2.	TASK METHODOLOGY	2-2
3.	CRITERIA FOR THE EVALUATION OF SERVICE AVAILABILITY MEASURES	2-6
4.	GENERAL COMMENTS REGARDING SERVICE AVAILABILITY	2-9
	APPENDIX	
	SUMMARY OF FIELD CONTACTS	2-13

LIST OF FIGURES

Figure 2-1. Diagram Illustrating Relationship of Transit Industry Groupings to Phases of Service Availability Life Cycle . . 2-4

1. INTRODUCTION

For some time, there has been considerable concern over the definition, measurement, and specification of a transportation system's effectiveness for providing service to its passengers in the face of the failure characteristics and consequences inherent in the design and operation of that system. Trip reliability, schedule adherence, compactness of trip time distribution, average delay, expected delays, availability, dependability, etc.: all are used at one time or another to describe this effectiveness measure. The net result is a welter of nonstandard terminology which serves not only to confuse the analyst but also to totally mask performance comparisons among alternative systems.

Accordingly, this study (a part of UMTA's Automatic Guideway Transit Technology program) is aimed at developing a set of measures for "service availability" which will be meaningful, readily understandable, and acceptable to transit operators, suppliers, and interested Government agencies. Service availability is defined in a generic sense as a measure of the impingement of equipment failures on the operation of a transit system as perceived by the users and operators.

Task 1 of this study consisted of an in-depth review of existing literature dealing directly or indirectly with the generic subject of service availability. Specifically, definitions, use, methods of measurement, models, and concepts as treated in the literature were sought. The results of this effort were reported in the First Interim Report.* The purpose of Task 2 was to carry out this information-gathering activity to the transit industry

* First Interim Report on The Development of Measures of Service Availability. Task 1.

to gain the benefit of their experience in the use of service availability measures. Specifically, Task 2 was directed toward gathering "real-world" insight into

- Service availability concepts/definitions
- Use of service availability measures at various phases of a transit system life cycle
- Factors influencing service availability and its use
- Characteristics of a "good" measure of service availability.

One major output of this task was a set of criteria to be applied in selecting/developing service availability measures for further processing in this program.

2. TASK METHODOLOGY

There are numerous types of parties to the development and operation of a transit system. Each of these parties deals with service availability, explicitly or implicitly, in a different way and for different reasons. It was the intent of this task to engage representatives from each of these parties in a dialogue to ascertain insights into service availability as viewed from their particular perspective and orientation with respect to the transit system life cycle. The first step, therefore, was to partition the transit industry into groups which shared a common perspective regarding service availability.

This was accomplished by defining a "life cycle" for service availability and partitioning the transit industry according to the interaction of various groups with the elements of this life cycle. Six phases were defined for the service availability life cycle.

(1) Establishment of desired goal. In this phase, the desired level of service is established.*

(2) Establishment of system level requirements. In this phase, this goal is translated into system level specifications.

---

* More often than not, the service availability concern is more fictitious than fact, resulting in little more than lip service.

(3) Allocation of system requirements to subsystem/ component requirements.

(4) Establishment of inherent service availability characteristics in design/manufacture. In this phase, the requirements are translated into hardware form.

(5) Growth of service availability/compliance testing. In this phase, which corresponds to system tests and evaluation prior to acceptance, an iterative process of service availability measurement and system modification is undertaken. This is a "debugging" phase prior to revenue operations.

(6) Availability maintenance. This phase corresponds to revenue service where the concern is maintaining adequate level of service availability.

The transit industry is divided into five categories as follows:

(1) Planning Agencies/Planning Consultants. This group establishes the desired nature of the transit system.

(2) Design Consultants/System Procurement Managers. This group translates the general planning results into hardware system requirements. Depending on the particular situation, they may establish specifications, prepare contract documents, evaluate proposals, conduct negotiations, supervise construction and installation, and oversees initial tests and evaluations to determine compliance with requirements.

(3) System Suppliers. This group translates the specifications into working hardware.

(4) System Operators. This group is concerned with the revenue operations of the transit system.

(5) Government Agencies. Various agencies of the Federal and local Governments which participate in any of the above activities.

Figure 2-1 shows the relationship of these transit industry groupings to the various phases of the service availability "life cycle".

Within this framework, nine representatives from the industry were contacted and visited for in-depth discussions of service availability as viewed from their perspective. These contacts were heavily biased toward

Service Availability Life Cycle Phase \ Industry Group	Planning Agencies Planning Consultants	Design Consultants	Suppliers	Operators	Government
1. Goal Establishment	●			●	●
2. System Requirements		●		●	●
3. Subsystem Requirements		●	●		●
4. Design/Maintenance			●		
5. Testing		●	●	●	●
6. Maintenance				●	●

FIGURE 2-1. DIAGRAM ILLUSTRATING RELATIONSHIP OF TRANSIT INDUSTRY GROUPINGS TO PHASES OF SERVICE AVAILABILITY LIFE CYCLE

automated system suppliers and operators. (The Appendix lists the companies/interests contacted.) The literature review (Task 1) reflected to a considerable degree the concerns of the planning community. One contact was made with a design consultant. The information gained is supplemented by Battelle's experience in the specification/monitoring/testing practices of new transit system installation. Government concerns are considered to be reflected in past programs (e.g., Dual Mode, HPPRT) and current programs (e.g., AGRT). Additionally, interaction with UMTA on the Downtown People-Mover Project is providing real data on the perspective of the Federal Government with respect to service availability.

In general, the interview approach was to engage the parties in a dialogue regarding the following:

(1) Background of service availability thinking in their specific areas of concern

(2) Definition of service availability

(3) Role of service availability in various phases of a transit system life cycle

(4) Problem of specification, measurement, and monitoring of service availability

(5) Desired characteristics of a good service availability measure.

With this information, a set of criteria to be used in evaluating alternative measures of service availability were deduced. These are contained in Section 3.0. Section 4.0 contains observations on the concept and use of system level service availability measures which are germane to the program.

3. CRITERIA FOR THE EVALUATION OF SERVICE AVAILABILITY MEASURES

At the outset of this study, several characteristics were hypothesized for a "good" measure of service availability. The Request for Proposal for this contract also stated several desirable characteristics. The field interviews offered no surprises nor, in retrospect, should they have. Basically, 10 criteria have evolved for evaluating service availability measures.

(1) It should reflect passenger perception of service.

(2) It should reflect the performance of elements of a transit system over which the operator can exert control.

(3) It should be capable of measuring improvement in performance due to operator control action.

(4) It should lead to clear, unambiguous performance specifications which can be directly treated in the design and manufacture of a transit system. By the same token, it must be measurable with reasonable data collection efforts, free of ambiguity, during tests and revenue service.

(5) It should reflect system performance goals rather than subsystem requirements so as not to unnecessarily constrain the suppliers.

(6) Like (2) above, it should relate to the elements over which design control can be exerted (nominally, the failure characteristics of and restore strategies of the technology being proposed and/or constructed).

(7) It should be sensitive to small changes in design and/or operational parameters.

(8) It should have technical validity.

(9) It should be independent of any specific technology. If specifications are the result of allocating system level goals, this allocation should _not_ unknowingly favor one technology over another.

(10) It _must_ be an effective communication tool among all elements of the transit industry.

As indicated above, these are not unusual or surprising. They are applicable to <u>any</u> performance measure and reflect rather simplistic but sound engineering judgments. While establishing these was a primary objective of the field interviews, it was, in fact, a very secondary result. Of far more significance to this project is the apparent "oneness" of purpose among all groups interviewed with respect to the concept and use of system level service availability measures.

> There is no controversy over the ultimate service availability measures desired. All parties subscribe to the concept of reflecting and controlling passenger <u>delays</u>, either in terms of frequency and/or duration.
>
> There was no controversy over the use of a system level measure for specification and/or performance monitoring. It was aptly pointed out that such system thinking has always (in recent past anyway) been a driving motivation behind specifications. What was missing is the explicit treatment. Implicitly, however, specifications were believed to accurately represent the "intent" of minimizing delay occurrence and duration. Operators realize that allocation of scarce resources during revenue service requires a measure of performance degradation and its relationship to causal factors--end-point items which can be attacked and controlled. Again, the performance degradation is considered to involve passenger delay of one form or another. Since this is difficult, if not impossible to measure directly, various proxy measures are employed, such as headway variations, schedule adherence, etc. While not directly measuring delay performance, these measures directly relate to delay performance. Improvements in schedule adherence, e.g., will result in improved service.
>
> There is no controversy over the <u>value</u> to be placed on a measure for specification purposes. Everyone agrees that, at this point in time, if a system level performance measure is used, no one has the market intelligence or passenger utility insights to establish the value. The supply community is obviously concerned because of the potential for unreasonable requirements. We have found, however, that all groups are equally concerned and attempting to act responsibly and equitably.
>
> There is no controversy over the fact that service availability, by itself, is insufficient to define the failure propensity and restore requirements of a transit

system. Allowable maintenance costs/capabilities are also extremely significant and, in the end result, may limit design decisions to enhance service availability by itself.

There is no controversy over the real problem to be faced in using system level service availability measures. This problem does not involve the "what" and the "why" of a measure but, rather, the "how". This is discussed in the following paragraphs.

As alluded to above, it is generally conceded that system level measures reflecting passenger delay propensity and properly accounting for the system variables which can be controlled is a desirable, in fact, a very desirable goal. With all this agreement, what is the problem? Discussion of this topic consumed the major part of our field interviews. Basically, the problem revolves around a series of "hows"--all interrelated.

How can a delay measure or a reasonable proxy measure be specified?

AND,

How can it be related to specific requirements for specific technological alternatives?

AND,

How can a proposal indicate compliance with the specification?

AND,

How can it be measured in actual operation to determine compliance?

AND,

How can it be measured during revenue service to maintain service availability?

Obviously, what is at issue is not the concept or the specific measure involved but, rather, the methodology for using it. A methodology which satisfactorily answers all of the above questions and can be demonstrated to do so is what is required. Therefore, perhaps the most significant criteria to be applied to a measure of service availability is that it be such that a suitable application methodology exist. Some of the characteristics of a "suitable" methodology are as follows:

(1) It should be simple to apply.

(2) It should produce repeatable results.

(3) It should not be sensitive to the scale of the system.

Additionally, nearly every criteria of a good performance measure is applicable to a good methodology. The net result of these findings is that an assumption of the existence of a single measure, applicable to all systems at all phases of the life cycle is probably invalid--primarily because of methodology constraints and opportunities posed by different levels of system complexity, scale, and technological sophistication. It is not at all inconveivable that each new system may require a unique variation in service availability definition and application methodology. This is an important conclusion and will receive full attention in the remaining tasks of this program.

4. GENERAL COMMENTS REGARDING SERVICE AVAILABILITY

This section highlights certain findings relative to the concepts, use, and measurement of service availability measures as derived from the field contacts. Not all are uniquely relevant to the study--many are basic suggestions for general consideration in the development of new AGT systems.

(1) As was discussed in the previous section, there is general agreement on the concept of using passenger delay (frequency and/or duration) as the service parameter to be controlled. However, direct use of a delay criterion is problematical.

   (a) Existing transit systems do not have the data collection and processing capability to measure such performance.

   (b) It is possible that such direct measurement could be done with new, automated systems, providing adequate data gathering and processing capability is built in.

   (c) Direct delay specifications pose a severe problem of translation into system failures/restore performance requirements. As indicated in the Task 1 report, computer simulation of system response to failure appears necessary. Using simpler models poses questions of validity and

scaling fidelity. If full simulations are required to allocate reliability/maintainability requirements, such simulation is only useful in the context of a specific technological offering. This poses severe costs on proposers for new applications and proposal evaluators who must judge the validity of the simulation. (Perhaps generalized models such as are being developed by the SOS contractor could help with the latter problem but the cost element would still be present.)

(2) There is a feeling that vehicle-based measures would be a reasonable and more tractable approach to performance specification and measurement.

(3) Regardless of the specific measure selected, new AGT systems should have provisions for collecting and processing vehicle performance at a minimum and possibly passenger delay performance. This would provide for the accumulation of a performance data base upon which future service availability performance could be based. It is significant that most existing AGT systems have these data somewhere in the system software with no ready means to process it into a useful form. Additional software is required.

(4) There is general agreement that high service performance results from effective failure management systems to minimize system restore times following a failure. A significant part of low restoration time capability is the ability to quickly identify the real cause of the failure. New AGT systems should find it almost axiomatic that built-in diagnostic capability will be cost effective.

(5) Regardless of the precise form of the specifications for a new AGT system, the intent of the specifications should be clearly stated--perhaps with provision for considering exception to the specification if it can be shown that alternative values meet or exceed the intent.

(6) There is a growing "system" orientation in the transit industry. The groups contacted appear to be comfortable with the concept of system-level performance specifications.

(7) Allocation of the system-level performance requirements to subsystem requirements must be done by the system designer.

(8) All too often, transit system planning is performed under the assumption of no failures, resulting in network configuration, vehicle size, capacity, etc. which are intolerant of failures. Rather than conduct these planning activities under the assumption of success, they should be conducted under the assumption of failure. Fault-tolerant designs and high service availability must be integrated into the system from its inception. It cannot be an add-on item. It is of more than passing interest that many other transit system features, such as comfort, lighting, etc., could be treated as add-on items but are treated very early in the planning process.

(9) Creating the image of a fault-free transit system during planning and marshaling for bond-issue support may be politically expedient but it serves to stimulate over reaction when the real operating system experiences the failures it undoubtedly will.

(10) The myth of perfection must be abolished. Every transit system must undergo a period of debugging and improvement (a period of service availability growth). This must be recognized in the procurement phase with provisions for a significant supplier effort during this period.

As a final note, one of the most encouraging findings of the field interviews was the spirit of cooperation exhibited by each of the contacts. Each group contacted shared the desire for system-level performance specifications and each exhibited not only concern from his own viewpoint but also an appreciation for the viewpoint of other elements of the transit industry.

APPENDIX

SUMMARY OF FIELD CONTACTS

This appendix briefly summarizes some of the information gathered during the field contacts.

OTIS/TTD

Contact Date: February 28, 1977

Persons Involved: W. Womack, D. Wilson, D. Dreith

Otis is one of the three AGRT contractors for UMTA. They previously had been involved in the dual-mode studies and the HPPRT program. In all of these programs, they have worked with the concept of service availability. In the AGRT program, they are using delay frequency and duration criteria supplied by UMTA together with an availability assessment model developed by Frank Smith and supplied by UMTA. This model uses Monte Carlo simulations of a simple network to assess the delay potential for an average passenger, given major subsystem failure rate as service restore time.

They are using their own analytic interpretation of the model as a guide in allocating system requirements and subsystem requirements.

TTD would like specifications at a system level to afford them maximum flexibility to capitalize on the particular strengths of their technology. They similarly support the need for built-in performance measurement capabilities.

While the concept of system-level specifications was embraced, TTD expressed great concern (seconded by other contacts) that a more important issue is the value imposed. Because there are no current data on passenger delay performance of existing systems, there is no benchmark to establish a reasonable goal. An arbitrary setting of maximum allowable delay could lead to feasibility problems. They suggested that market research needs to be done to determine how the public views failure-induced delays and what their tolerances might be.

BOEING, KENT SPACE CENTER

Contact Date: March 1, 1977

Persons Involved: R. Tidball, V. Leiskow

Boeing is also an AGRT contractor using the service criteria and assessment models supplied by UMTA. Concern was expressed over the scaling fidelity of this simple model to a complex network. They expressed concern that one could design to look "good" in the simple model but be "bad" in an overall network situation. They are investigating this potential.

Boeing indicated that future AGT systems possess an inherent capability which is lacking in current systems, that being the capability to automatically collect and process performance data--even to the level of passenger delay. Any system-level performance specification should be accompanied by a requirement to automatically collect measurement data.

Note: It is apparent that OTIS/TTD, Boeing, and presumably Rohr, in the context of their AGRT contracts, are engaged in service availability application methodology. Their experience in this activity should go far in pinpointing the problems alluded to in Section 3.0 and possible solutions.

SEATTLE TACOMA AIRPORT, SATELLITE TRANSIT SYSTEM (STS)

Contact Date: March 1, 1977

Persons Involved: M. Bitts, J. Borkowski

Specifications for the STS were of three types.

- MTBR/MTTR/availability requirements on 21 systems
- Loop time of 5 minutes
- Service from central terminal to satellites every 2 minutes.

Of these, the latter was taken as the most crucial. It became apparent that adequate reliability of service to the satellites could not be achieved with MTTR's as specified (on the order of minutes to hours). Consequently, performance was improved primarily by decreasing MTTR. Currently, the average time to restore service is 3 minutes. (They define a failure as any service interruption which requires the dispatch of a service man. Time

to restore service is measured from the onset of the disruption to the time service is restored, either under automatic or manual control.)

The STS is in its "availability maintenance" phase, that is, its failure performance has reached a steady-state condition and patterned failures have been mostly eliminated (through design or preventive maintenance). Current efforts involve developing patterns of other failure modes to enable them to be eliminated. This involves increasing knowledge of precise failure causas factors. This same information is required to reduce downtime by providing rapid diagnostics for the service man. To provide the capability, STS is now building a Vehicle Data Acquisition System (VDAS).

BECHTEL CORPORATION

Contact Date: March 2, 1977
Person Involved: J. Williams

Bechtel has been active in the specification and design consultation of major rail transit systems. It was part of a consortion for BART and is currently working on MARTA and a new system for Caracas, Venezuela.

They have long been concerned about the faithfulness of specifications of the MTBR/MTTR type in representing the real service intent of the transit system. While, in the past, these were derived from passenger service considerations, this was not done in an explicit, rigorous manner. Bechtel is attempting to define a useful service availability specification and measurement procedure for use in the Caracas job. While not fully formulated yet, their initial thinking is to use train headway variation induced by failures as a measure. They would then allocate requirements to basic systems for contract specifications. They have informally communicated with potential bidders about the possible form of the specifications and have not experienced any resistance.

Note: This cycle of pre-RFP discussions regarding the form and wording of the specifications would appear to be very useful in circumventing interpretation questions and problems during proposal preparation. It should be considered whenever system-level specifications are used, where even the work "system" is misinterpreted.

BAY AREA RAPID TRANSIT (BART)

Contact Date: March 3, 1977

Person Involved: J. King

BART has an extensive data-gathering system on vehicle failures, maintenance actions, and status. This is a real-time system and is of great utility in assessing vehicle performance. BART is now using a trip schedule adherence measure (percent of end-to-end trips completed within 5 percent of scheduled run time) to measure performance. By correlating offsets with specific failures, the "bad actors" can be identified, aiding in the allocation of scarce maintenance/operating funds.

It was pointed out that, if at all possible, AGT systems should be designed, built, and operated as a series of separate systems, utilizing passenger transfers to achieve origin-destination variations. Such separation of function minimizes the overall system impact of a single vehicle failure. It was pointed out that the Oakland WYE is a constant trouble spot because of the system implications of a failure at that point.

It was emphatically suggested that "no-transfer" policies for new AGT systems be reevaluated to permit such separation of links. By allowing convenient transfers, the overall system performance may be enhanced. (This was also supported by most of the other contacts.)

DALLAS/FT. WORTH AIRPORT - AIRTRANS

Contact Date: March 4, 1977

Persons Involved: D. Ochsner, D. Elliot

The specifications for AIRTRANS were MTBF/MTTR at a major subsystem level. There was no overall service goal (from a fault-tolerant standpoint) nor a compliance demonstration procedure. A system simulation was performed to determine required schedule frequencies and reliabilities which were then measured for acceptance tests.

Currently, AIRTRANS has the capability of collecting vast quantities of performance data and is doing so. There are no provisions for processing these data into useful information, however.

The failure patterns experienced by AIRTRANS were analyzed and categorized into a few types. A subjective evaluation of the effects of these failures in terms of delay/inconvenience to the rider resulted in weighting factors which are used to compute a daily figure of merit. This method has been successful in tracking performance and reflecting the results corrective action.

VOUGHT CORPORATION

Contact Date: March 4, 1977

Persons Involved: D. Benjamin, R. Raven, D. Randolph, W. Pitts, C. Schultz, A. Songayllo

Like the other system suppliers contacted, Vought would like the use of a system-level performance specification, subject to its ability to be measured and predicted. Fairly lengthy discussions were held over application methodology. These served a useful function of highlighting the problems, if not the solutions.

Vought supported the thesis that the system supplier is not finished at system delivery. A lengthy period of debugging and availability growth ensues which must be accounted for in the procurement.

MORGANTOWN PRT

(Telephone discussion with Phil Morgan, UMTA)

The current performance measurement is service dependability which is the series product of system availability, fleet availability, and trip reliability. In the computation, only the link between Beechurst and Engineering is counted. (This is rationalized because of the low passenger volume on the Walnut Street link.) Thus, the "system" is a single link with a station at either end. The simplicity of this system definition gives meaning to the term "system availability". The system is considered available when service capability exists between these two stations and unavailable when service is interrupted.

In the expanded system, however, system availability becomes obscure. A failure in one link will not shut the entire system down. Service will still be available on other links. As a result, Boeing has proposed that system availability be defined as the weight sum of link availabilities where the weighting factor reflects the expected link trip volume to the overall system trip volume.

Other factors in the dependability calculation will remain the same. There will be a change, however, in the base against which fleet availability is computed. Currently, the base for this calculation is the number of vehicles required to meet capacity demands. These are defined from a table and reflect the demands expected during initial system planning. The new base will be the number of vehicles procured with a target goal of having 85 percent of these in operable condition at any point in time.

PART 3

THE DEVELOPMENT OF MEASURES OF SERVICE AVAILABILITY

TASKS 3 AND 4. SERVICE AVAILABILITY MEASURES DEVELOPMENT AND GUIDELINE DOCUMENT

CONTRACT NO. DOT-TSC-1283

to

DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER

by

R. D. Leis

BATTELLE
Columbus Laboratories
505 King Avenue
Columbus OH 43201

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	3-1
	1.1 Task Objectives	3-2
	1.2 Task Procedure	3-3
	1.3 Organization of the Document	3-4
2.	RESEARCH FINDINGS	3-5
3.	THE SERVICE AVAILABILITY CONTROL PROBLEM	3-7
	3.1 Service Availability--User Perspective	3-7
	3.2 The Conceptual Service Availability Control Process	3-11
	3.3 Assessment of Existing Service Availability Measures	3-16
	3.4 Development of a Methodology for Relating Passenger Delay Potential and System Failure Characteristics	3-20
4.	SERVICE AVAILABILITY CONTROL PROCESS	3-21
5.	METHODOLOGY FOR RELATING SYSTEM FAILURE CHARACTERISTICS AND PASSENGER DELAY CRITERIA	3-21
6.	AN APPLICATION EXAMPLE	3-21
	APPENDIX A	
	ASSESSMENT OF EXISTING SERVICE AVAILABILITY MEASURES	3-23
	APPENDIX B	
	METHODS FOR ESTIMATING FAILURE-INDUCED DELAY PARAMETERS	3-47

LIST OF FIGURES

Figure 3-1.	Alternate Theories of Passenger Perception of Trip Delays	3-9
Figure 3-2.	Conceptual Service Availability Control Process	3-12

LIST OF ABBREVIATIONS, SYMBOLS, AND SPECIAL TERMS

α - Dummy variable computed to simplify station delay calculation in short-cut method

λ - Failure rate, number of failures divided by operating time

AGT - Automated guideway transit

Buyer - General term designating procurement and operating agency, including consultants in support of such agencies

CH - Cushion Headway, difference between normal time spacing of vehicles and minimum spacing allowed by safety and/or operating constraints

C_v - Vehicle capacity

D - Cumulative delay incurred by delayed passengers due to a failure or combination of failures

D* - Cumulative delay incurred by delayed passengers due to a unit failure

$\overline{D}$ - Average delay per delayed passenger

$\overline{D}_v$ - Average delay per delayed vehicle

DPM - Downtown people mover

DR - Total system demand rate. Passenger trips per unit time

EC_a - Average excess capacity available at any given station to dissipate queues subsequent to a failure

EC_n - Normal excess capacity. Available system capacity in excess of that required to meet passenger demands during normal (unfailed) operation

ED - Expected delay, average delay on an average trip

FMDEA - Failure mode and delay effect analysis

LF - Load factor. Measure of vehicle utilization equal to number of passengers on board divided by vehicle capacity

LOS - Level of service

LR - Link flow rate, passengers per unit time traversing link

MTBF - (Mean-Time-Between Failure). Measure of failure frequency equal to the operating time divided by number of failures observed

MTTR - (Mean-Time-To-Restore). Measure of maintainability, in this report, MTTR is the mean duration of a failure

N_f - Number of vehicles operable during a failure

N_n - Number of vehicles normally in operation

N_r - Number of vehicles operating during service restoration subsequent to a failure. Applicable if excess capacity is derived from inserting extra vehicles into the system

N_s - Number of vehicles stopped due to failure in system which permit headway closure

PD - Number of passengers delayed due to a failure or combination of failures

PD* - Number of passengers delayed due to a unit failure

Pr_d - Probability of delay, likelihood of experiencing a failure induced delay on an average trip

RFP - Request for proposal

SAM - Service availability measure

Service Availability - Impingement of system failures on transportation service as perceived by passenger

SRT - Service restore time. Time interval following TTR required to dissipate station queues to normal values

Supplier - General term designating manufacturers, contractors, consultants, etc., engaged in design, construction, manufacture of AGT systems

TT - En route time for average trip

TTR - Time to restore for a specific failure

$\overline{TTR}$ - Mean time to restore

$\widetilde{TTR}$ - Quadratic mean time to restore (RMS value of all TTR values)

V_f - Vehicle velocity during a failure

V_n - Normal vehicle velocity (average)

V_r - Vehicle velocity during service restoration. Applicable if excess capacity is derived from increasing vehicle speed (average)

1. INTRODUCTION

For some time, there has been considerable concern over the definition, measurement, and specification of a transportation system's effectiveness for providing service to its passengers in the face of the failure characteristics and consequences inherent in the design and operation of that system. Trip reliability, schedule adherence, compactness of trip-time distribution, average delay, expected delay, equipment availability, dependability, etc.; all have been and are being used in one form or another to describe this effectiveness measure. No standard approaches exist; no standard terminology exists; and no standard methodological framework exists for establishing performance goals and controlling system design and operational parameters pursuant to these goals.

Accordingly, this study (a part of UMTA's Automated Guideway Transit Technology program) is aimed at developing a set of measures for "service availability" which will be meaningful, understandable, and acceptable to transit operators, suppliers, and interest Government agencies. Service availability is defined in a generic sense as a measure of the impingement of equipment failures on the operation of a transit system as perceived by the system users and operators.

Task 1 of this study consisted of an in-depth review of existing literature dealing directly or indirectly with the generic subject of service availability. Specifically sought were definitions, use, methods of measurement, models, and concepts as treated in the literature. The results of this effort were reported in the First Interim Report. Task 2 carried this information-gathering activity to the transit industry to gain the benefit of their experience in the use of service availability measures. Specifically sought were "real-world" insight into

- Service availability concepts/definitions
- Use of service availability measures in various phases of a transit system's life cycle
- Factors influencing service availability and its use
- Characteristics of a "good" measure of service availability
- Criteria by which alternative measures can be evaluated.

The results of this activity were reported in the Second Interim Report.

The original program schedule called for the sequential performance of a task of selection of appropriate service availability measures and a task to develop and demonstrate the methodology for utilizing those measures. An important conclusion of Task 2 was that such separation cannot be made. Perhaps the most important criterion for a good measure is the existence of a simple, understandable, and usable methodology for its use. Hence, these tasks were combined and their results are reported in this document.

1.1 Task Objectives

The objectives of Tasks 3 and 4 were to

(1) Select an appropriate measure or measures of service availability according to the criteria defined in Task 2, and

(2) To demonstrate the use of such a measure or measures via a handbook-like document.

During Task 2, several insights were gained which complicated the anticipated procedures for reaching these objectives. Additionally, and more importantly, the discussions with field representatives caused real concern over the likelihood of success of achieving these objectives. Briefly summarizing, all field representatives, regardless of their positions relative to the various phases of a transit system's life cycle, felt that the only performance measure of concern was one which treated passenger delay potential--either frequency, duration, or both. Equally emphatic was the belief that utilizing such a measure was as problematical as to become impractical. Therefore, what is desired is a hardware performance measure,

with a direct relationship with the parameters designers and operators can manipulate and which is a faithful representation or proxy for passenger delay propensity.

It was recognized that numerous measures and/or models exist which claim to accomplish this. In this regard, concern was expressed with respect to both the validity of these models and their utility in a real-world service availability control process.

1.2 Task Procedure

To respond to the concerns expressed by the transit industry, the initial activities of the task shifted from the relatively simple procedure of evaluating existing measures of service availability to a more fundamental investigation of the service availability control process. This was required for several reasons.

(1) To gain a full appreciation of the concern expressed by the transit industry

(2) To define the service availability control points and the role of service availability measures (SAMs) at these points

(3) To determine the real problem focus with respect to SAMs and their utilization

(4) To direct the derivation of a responsive solution.

A pragmatic inductive approach was selected. Simple transportation systems were assumed to permit deterministic manual simulation to define the parameters which influence passenger delay. By varying these parameters and assessing their impact on delay propensity, correlations were detected between passenger delay characteristics and system failure characteristics. Various hypotheses were developed and tested in other more complicated, assumed systems to determine validity, limits of applicability, and degree of generality. Throughout this process, various postures were adopted to represent the specification of performance as might occur in the system planning/procurement phase, the translation of these specifications into hardware requirements into performance monitoring options during the system operational phase.

It should be emphasized that while the example systems utilized were assumed, they were not taken to be unrealistic. In general, the examples consisted of loops, shuttles, and parallel guideways. Trips were assigned to these systems along fixed routes. Failure consequences of full and partial shutdown were considered. It is believed that the systems utilized exemplify most systems in existence today. They do not represent the network concepts which operate entirely in a probabilistic manner in terms of route selection and vehicle dispatch options. For these systems, assessing normal operations requires computer simulation. It is clear that predicting service availability will similarly require sophisticated simulation approaches. While the approach described in the report employs manual modeling and analysis, it is obvious that a computer could be used to do the same tasks.

1.3 Organization of the Document

The major findings of these task activities are presented in Section 2.0. Section 3.0 discusses the basic task activities, and provides the rationale and derivation of the concept presented in Sections 4.0 and 5.0. Two appendices support Section 3.0. Sections 4.0 through 6.0, as a set, constitute a guideline manual for controlling service availability of AGT system. Section 4.0 describes the control process; 5.0 describes the methodology for relating passenger delay potential and system failure characteristics as required by the control process; and Section 6.0 presents an in-depth example of the principles of 4.0 and 5.0. The reader who is familiar with these guidelines who wishes to explore the derivation will need to refer only to Section 3.0 and the two appendices.

2. RESEARCH FINDINGS

While the objectives of this research may have inferred the existence of a few selected measures which could be universally applied to transit systems, thereby standardizing the specification and control of transit system failure characteristics, this was by no means presumed. The general result of this research is that such a measure does not exist except at the passenger perception level itself. Translation of criteria imposed at this level into design and operating controls becomes unique to the system of concern--these must be derived for each application. Recognizing this shifts the problem emphasis from particular alternative measures to one of procedures for their development and use. Approaches for handling this latter problem in the form of guidelines constitutes the major result of this research. These guidelines constitute Sections 4.0 through 6.0 of the report.

The following specific findings support these guidelines:

(1) While it is generally conceded that passenger perception of the impact of off-normal performance on transportation service is not adequately understood, there is general agreement within the transit industry that this perception is influenced by two parameters.

(a) The frequency of delay events experienced by the passenger

(b) The duration of these events, individually and collectively.

Therefore, controls imposed on the failure characteristics of a transit system should be related to values imposed on these two parameters. Similarly, measurements of operating performance should be relatable to these parameters.

(2) Over the life cycle of a transit system, three generic service availability measures (SAMs) are utilized.

(a) The first SAM deals directly with the delay parameters in terms of allowable values for individual passenger exposure to delays in terms of frequency and duration. This is termed SAM 1 and constitutes the basic design and operational criteria.

(b) The second translates SAM 1 into allowable failure characteristics of a specific system alternative to be used to control the design/manufacture/delivery of the transit system. This is termed SAM 2 in this report.

(c) The last SAM relates SAM 1 and SAM 2 to off-normal performance characteristics which can be measured for controlling service availability during the operating life of the system.

(3) The propensity of a failure to induce delays is dependent on the specific nature of the transit system being analyzed. This relationship is complex and influenced by many system and application-specific variables.

(4) Because this relationship between failure characteristics and delay results is system specific, measures for controlling these failure characteristics must be system specific and must be derived for each system application. Thus, there is no general relationship or proxy measure which can be considered applicable across different transit systems. Such comparisons can only be made in terms of SAM 1 and these are only approximate because SAM 1, itself, is application specific.

(5) The relationship between system failure characteristics, off-normal performance, and delay parameters requires an acute understanding of system failure dynamics and the corresponding passenger dynamics. In a general sense, such an understanding can only be developed through system simulation techniques. For systems which are characterized by randomness, in route, schedule, origin-destination pattern, etc.; computer simulation will be required. For simpler systems, however, useful approximations of expected performance can be developed by manual techniques. These are dealt with in this report.

3. THE SERVICE AVAILABILITY CONTROL PROBLEM

The intent of the transit industry is to facilitate the transportation of people by providing services which are attractive to potential passengers. Within the context of service availability, this implies a provision of service sufficiently free of failure-induced interruption such that passenger perceived service does not degrade below some acceptable level. Having established parameters and values which describe acceptable levels of passenger perceived service, it is the role of the transit industry to control these parameters. It is the role of service availability measures to direct these control actions and to gage their effectiveness.

3.1 Service Availability--User Perspective

As defined in Section 1.0, service availability in its generic form is a measure of the impingement of failures on the transportation service provided by a transit system as viewed by the users of that system. In other words, service availability, as a measure of transit system performance, seeks to relate the failure characteristics of a transit system to some service parameter or group of parameters to which a passenger associates with service goodness (or badness). From the system operating standpoint, system failures evidence themselves as temporary losses in ability to transport passengers in the normal manner. To the passenger, this off-normal performance fosters a lack of confidence in the system's ability to perform in the "normal" manner; a lack of service predictability which is necessary to comfortably utilize the services provided.

The actual parameters of service degradation to which a passenger relates have not been absolutely defined. A passenger's perception of reliable, predictable service may take many forms, examples are:

Reliability of destination achievement
Waiting time
In-transit delays
Missed connections
Seat availability
Trip-time variance

To this list might be added effects of failure such as perceived safety, congestion, comfort, and other parameters which are equally nebulous. Literally, pages could be filled with alternate expressions which have been used at one time or another to describe the passenger perceived impact of system failures. The only useful point which would be served, however, would be to further underscore the fact that passenger attitudes with respect to system failures and their effects are not completely understood.

In view of the fact that the transit community (including planners, designers, A/E firms, operators, etc.) does not fully comprehend how the passenger "measures" the impact of failures; and in view of the fact that such knowledge is necessary to establish the dimensions by which failure characteristics must be measured, it becomes necessary to assume this passenger measure. In this regard, the concensus opinion of the transit industry is that passengers respond negatively to trip delays--in terms of both frequency and duration. While this is an assumption, it is quite reasonable and affords the translation of many attitudes into theoretically measurable quantities. Accepting the thesis, however, does not imply a concensus of how these parameters are viewed in a passenger's assessment of performance. Simply, the levels of delay frequency and/or duration which are considered to constitute bad performance do not enjoy a concensus opinion within the transit industry. Figure 3-1 illustrates examples of the types of theories proposed.

In general, the curves depicted in this figure illustrate an increasing level of annoyance as the delay duration increases. Curve I illustrates a linear relationship which implies that a passenger views a single delay of two units exactly as he would two delays of one unit each.

Curve II illustrates the theory that passengers are insensitive to "short" delays. Hence, one is concerned only with delay events which are longer than some predetermined "tolerable" level. This theory sounds plausible, but introduces the problem of defining a "tolerable" delay. While there is agreement that the tolerable delay is a function of trip parameters such as trip type, purpose, and normal travel time, there is no general agreement as to the value to be applied.

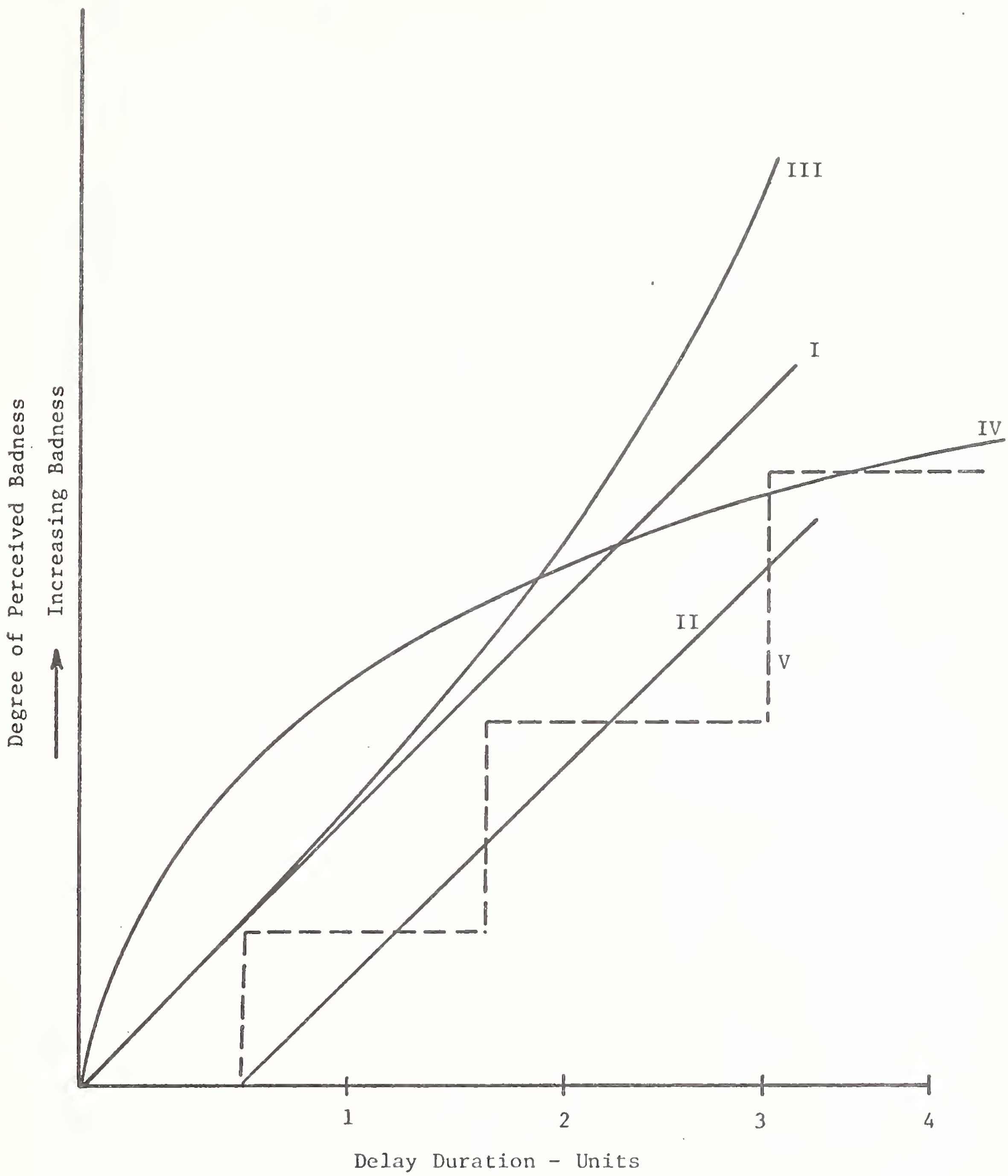


FIGURE 3-1. ALTERNATE THEORIES OF PASSENGER PERCEPTION OF TRIP DELAYS

Curve III represents the theory that a passenger's annoyance factor increases at a rate greater than the increase in delay duration. A delay of two units is more than twice as bad as a delay of one unit. Using this theory, one has to weight faulure occurrences by their duration to gain an overall performance measure. This theory is also plausible. We can all relate to instances where this type of response was present.

Curve IV illustrates an opposite viewpoint--a theory which says a delay event causes an initial surge of annoyance, the rate of which diminishes as the delay duration decreases. Like the previous theories, this too is plausible. A passenger subjected to a delay initially does not know if it will be long or short and his annoyance sensors may always be tuned to expect the worst. Again, we can relate to the feeling from our own experience.

Curve V represents a theory which can be like any of the previous ones except that it assumes that passengers do not evaluate delays in a continuous fashion. Rather, they are insensitive to delay duration over selected ranges. For example, a three-minute delay might be viewed with the same annoyance as a five-minute delay; or a twenty-five-minute delay may be considered equivalent to a thirty-minute delay. Again, this is a plausible theory; one which we can support as the basis of our own experience.

From these theories, we can see that the limits of tolerable delay can take many forms and values. In addition to these variations, the location and type of delay are also considered influential in a passenger's perception of service. Typically, passengers can experience delays induced by system failures in one of two ways.

(1) By being aboard a vehicle which is stopped or slowed due to a failure

(2) By being denied normal access to the transit system at a station due to a failure.

If one considers that passengers perceive only the excess trip time resulting from delays, the above distinctions are not significant. If, however, stations are exposed to severe weather conditions, one might well "weight" station delays higher than en-route delays.

Alternatively, station delays afford the passenger an option of aborting the trip, an option which does not exist if the passengers are

delayed en route. Hence, one might choose to attach a premium to these latter delays. Significance may also be attached to the type of en-route delay encountered. Failures which require passengers to stop may be perceived different from those which merely reduce vehicle velocity--particularly if the system uses tunnels or elevated structures.

Hence, it can be seen that "delay", as a measure of passenger perceived service, can take many forms and values. This consideration drives the necessary conclusion that a service availability control process must be capable of functioning within these variations.

3.2 The Conceptual Service Availability Control Process

Figure 3-2 illustrates the conceptual process of controlling service availability, beginning with the delay theories discussed in Section 3.1 and concluding with an operating system. As indicated in this figure, passenger delay parameters are not controllable in a direct sense. They can only be controlled indirectly by manipulating certain system design and operating variables. These are termed "primary" controls and the exercise of these controls is highlighted by a primary control envelope in Figure 3-2.

During the preoperational phase of a transit system's life, the signal driving these primary control actions is derived by comparing the predicted off-normal performance of a system design alternative, with its expected failure characteristics and assumed operating characteristics, with criteria established by SAM 1 through the transfer function f_2. This transfer function has special significance--this being its role in relating system off-normal performance and passenger delays induced by the off-normal performance. Hence, with an input consisting of allowable delay characteristics, exercise of this function results in allowable off-normal performance characteristics which form the hardware-oriented criteria for system design.

In terms of service availability measures, the input criteria to f_2 is termed SAM 1. SAM 1 is itself derived from interpretation of good service as perceived by the passenger. This SAM has units of delay; e.g., frequency and/or duration, and values consistent with the level of control desired. The output of f_2 is termed SAM 2. This measure is in terms of

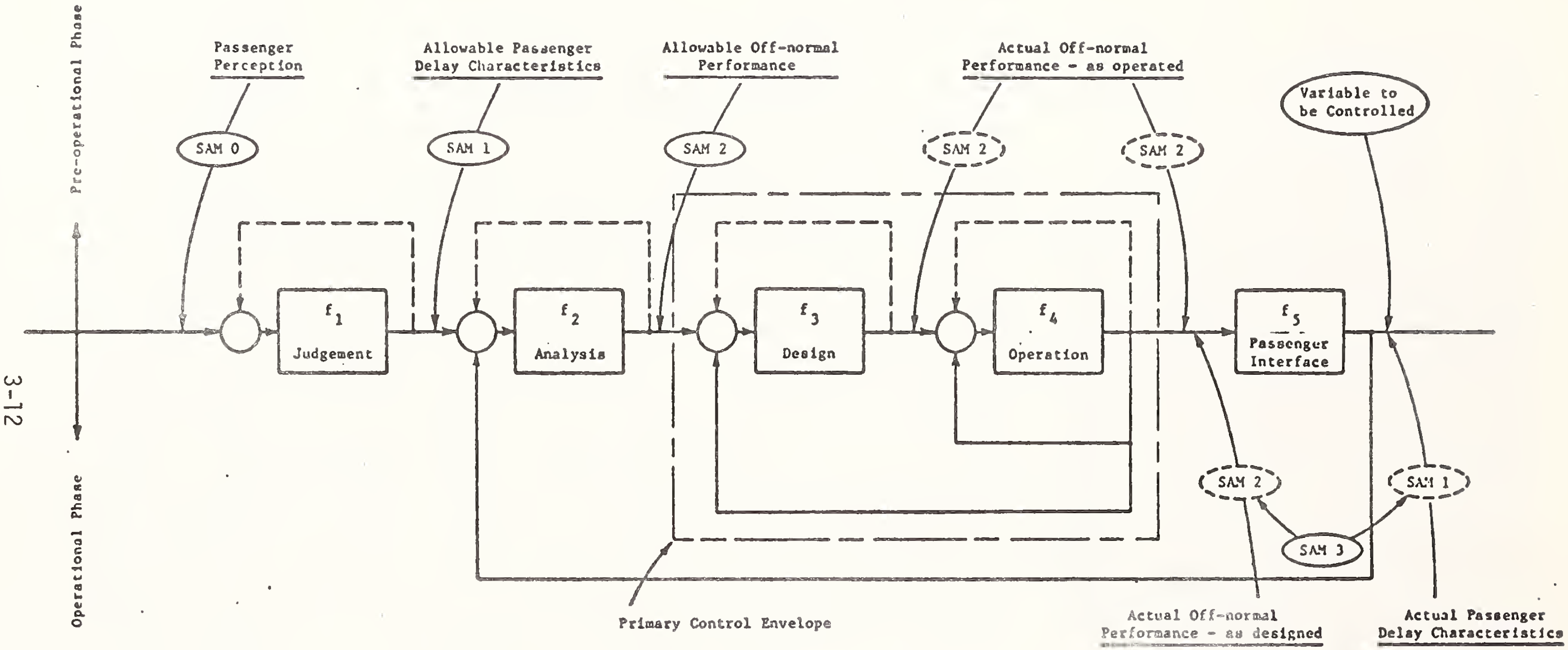


FIGURE 3-2. CONCEPTUAL SERVICE AVAILABILITY CONTROL PROCESS

system off-normal performance variables with allowable values consistent with the level of control established by SAM 1. During system design and manufacture (f_3) and the tests and evaluations performed on the finished system, SAM 2 is the reference point for acceptable performance.

During the operating phase of the transit system's life, the primary controls are still system design and operation oriented. However, the emphasis is placed on manipulating operating variables (f_4). The output of f_4 is the actual off-normal performance of the operating system. SAM 2 would be the reference point. The transfer function f_5 represents the interaction of this off-normal performance with the actual passenger trip demands and origin-destination patterns, resulting in actual passenger delay characteristics; hence, SAM 1 would be the point of reference for acceptable performance.

No standard scale exists to properly assess all of these possible situations to <u>derive</u> the appropriate treatment of delays in any given situation. Therefore, establishing the precise delay definition and allowable values for service availability control is a judgment call by the system buyer--consistent with his desires and goals relative to his specific application. This is not a trivial determination, however.

System cost can be influenced in a significant way by the precise delay parameters and values established as SAM 1. Hence, in establishing these criteria, sufficient insight into the implications for design, operation, and ultimately cost. Such insights may well serve to modify the criteria to a more realistic and acceptable value.

The formulation of SAM 1 can involve as many variables and theories as discussed in Section 3.1. For example, SAM 1 may be expressed as

Probability of incurring a delay on an average trip

Probability of incurring a station delay

Probability of incurring an en route delay

Probability of incurring an en route stoppage

Average delay associated with the above delays

Average delay encountered on average trip

Cumulative number of delays expected for an average passenger over some period of time

Cumulative delay experienced by an average passenger over some period of time

Exclusive combination of the above.

Because of its role as the control signal to drive the design of the transit system, SAM 2 must reflect the variables which can be controlled in the process: failure types, failure frequency, failure duration, etc. SAM 2 establishes controls for all system off-normal performance parameters which influence passenger delays. In this context, the analysis function, f_2 in Figure 3-2, takes on a special role; that of quantitatively relating the off-normal performance parameters to passenger delays to establish the appropriate control levels to be imposed. As will be discussed in later sections, the methodology to do this is a key element in not only establishing SAM 2, but also exercising the entire control process.

During system operation, controls are exerted on the operational aspects of the system (e.g., maintenance, failure management, etc.) and on design variables (e.g. retrofit procedures) to maintain or enhance actual delivered performance. As indicated in Figure 3-2, the feedback information to direct appropriate operational control actions can be obtained by SAM 3 which can take two forms.

(1) Off-normal performance measurements in the form of SAM 2. Measures of this form can be compared with the criteria of SAM 2 (established during the pre-operational phase) provide the necessary error signal to drive appropriate control actions. These actions may involve changes in operating procedures (e.g., changing maintenance procedures) or design (e.g., retrofitting failure diagnositic equipment or remote resettable failure correction). The assumption with this approach is that SAM 2, as previously established, represents "good" performance and that actual off-normal performance consistent with SAM 2 will result in acceptable passenger delay performance. However, SAM 2, of necessity, is derived based on certain assumptions. A particularly significant assumption (as will be discussed later in this report) involves the passenger travel demands and the origin-destination mix anticipated. Hence, SAM 2, as derived to control the design process, represents good expected passenger

delay performance in actual operation only if the actual passenger demands match those assumed during the exercise of f_2.

(2) Actual passenger delays directly in the form of SAM 1 for comparison to the criteria established by SAM 1. Such a measure circumvents the assumption problems of (1) above. However, to translate "bad" performance into appropriate control actions, the relationship between delays and system off-normal performance must be established. If this relationship is assumed to be f_2, as derived during the design phase, the same correlation problems exist. Furthermore, measuring actual passenger delays is a problem. In new AGT systems, mechanisms could be incorporated into station equipment to measure passengers delayed en route. Measurement of passenger delays at stations are, however, difficult and, if possible at all, would require special techniques.

Alternatives exist to circumvent problems arising from both of the above measures. First, the actual transport demands experienced in operation could be ascertained and a "new" SAM 2 could be derived as a standard for comparison. Alternatively, measures could be taken of the off-normal performance and compared with periodic measures of passenger delays (by survey or sampling techniques, e.g.) to establish a relationship between passenger delay and system off-normal performance. In either case, the net result is a recalibration of SAM 2 for use in evaluating system performance.

In summary, exercising control of service availability over the life of a transit system requires translations of passenger-perceived service characteristics into a meaningful form to interface system design characteristics and operating variables--the two system aspects which can be controlled. While, in any specific transit system evolution, many translations and SAMs may be used, generically three are required.

(1) SAM 1 - This represents the buyers judgmental interpretation of passenger-perceived attributes of good service. It consists of delay parameters and is the general performance goals from which other SAMs are derived.

(2) SAM 2 - This measure is an equipment-related measure, defining the units of off-normal performance and associated values as derived from the criteria established by SAM 1. SAM 2 is the criteria by which system performance in an as-designed condition will be evaluated. The form of SAM 2 is variable, being a function of

the particular off-normal performance characteristics of a particular transit system. In general, however, SAMs in this category should be functions of system failures and restore characteristics.

(3) SAM 3 - This measure may be in the form of SAM 1 or SAM 2, depending on data-collection abilities. More important than the specific form is the interpretation of the values measured. Determination of "good" performance may be assumed on the basis of comparison to the criteria of SAM 1 or SAM 2 if the system characteristics assumed during the preoperational phase are replicated in actual operation. If not, the operator must establish new criteria, but need not redefine the model used.

Figure 3-2 was constructed to be logical representation of the control actions required to effectively control service availability throughout the life cycle of a transit system. The thrust of the figure, as well as the previous discussion, has been on control signals, not control actions. It is assumed that various parties exercising a particular control can manage the resources at their disposal if the appropriate criteria of evaluation are made available to them. While the control process shown in Figure 3-2 is represented as a conceptual framework, it is not considered to be only theoretical. Quite the opposite, each step depicted must be carried out. Indeed, these have been carried out in the past although, in many instances, not explicitly. A major conceptual aspect of Figure 3-2 is the explicit recognition that SAMs can be derived from one another and, if done properly, can replace one another, as appropriate, to the particular control action being taken. The key element in this process is the derivation of SAM 2. SAM 2 must be derived from a proper comprehension of the failure characteristics of the system and the delay impacts of these failure characteristics. It is understandable that most of the effort in the service availability area has been directed toward defining SAM 2 and the analytic models for deriving SAM 2. In this program, this also emerges as the problem focus.

3.3 Assessment of Existing Service Availability Measures

During Task 1 of this program, numerous measures and analytic models for assessing service availability were identified. These measures,

as indicated by their form, consist of system failure characteristics, combined in such a way as to compute a value which is <u>interpreted</u> in terms of some passenger delay parameter (generally a probability formulation). Ostensibly, they fulfill the requirement of relating system off-normal performance and passenger delay. The project intent was to define a set of criteria (during Task 2) by which these measures could be evaluated--to judge their applicability and utility as effective service availability control parameters. However, all of the criteria generated address, implicitly or explicitly, two major underlying concerns.

(1) Do these models and measures really represent the relationship between passenger delays and system off-normal performance?

(2) Do these models and measures lend themselves to reasonable and understandable techniques for analytic determination and/or measurement?

Therefore, to properly assess existing SAMs, it was necessary to derive an understanding of the dynamics of failure-induced passenger delays in simple example transport situations where passenger delays could be easily visualized. This understanding provides certain "baseline data" which should be replicated by a valid service availability measure. Furthermore, and more importantly, in the eventuality that no existing measures survived this test, this understanding provides the basis for taking remedial action.

Appendix A discusses this activity in detail. The important findings of this activity are discussed in the following paragraphs.

The delay environment to which a passenger is exposed can be illustrated by a simple scenario. During normal operations, passengers enter stations in some random fashion and are transported to their destination according to the service schedule of the system. This demand and supply process constitutes the expected performance as viewed by the passenger. If the systems were to experience a complete shutdown, all passengers en route at the time of the failure would be stopped for the duration of the failure. Additionally, all passengers entering the stations would be denied service. At a minimum, this service denial would exist for the duration of the failure. However, even after the equipment elements of the system have been restored to normal operating conditions, passenger delays may continue

to accrue at stations until the delayed passenger queues are completely dissipated. Hence, station delays accrue during the time to restore equipment to normal operation (TTR) and during the time required to dissipate the queues which developed during TTR. This latter time is termed SRT to denote "service restore time".

By generalizing this simple scenario, as developed in Appendix A, the system parameters which influence passenger delays were determined to be:

(1) Failure type - classified by the effect of the failure on the ability of the system to deliver required capacity in the vicinity of the failure. Three general types are considered: (a) failures which result in a blockage, (b) failures which result in operations at velocities less than the normal velocity, and (c) failures which result in operations with less than the required number of vehicles.

(2) Failure rate - the expected number of failures in some unit of time.

(3) Failure duration - the time during which the failed state exists.

(4) Failure location - the location of the failure relative to the general system configuration. This is important where failure tolerance is provided.

(5) System failure tolerance - the ability of the system to limit the impact of certain failure situations by bypassing or otherwise disconnecting the failure affected area. This feature determines the extent to which a specific failure disturbs total system performance.

(6) Passenger trip demands - in terms of the quantity of trips requests per unit time.

(7) Trip origin-destination patterns.

(8) System capacity - more appropriately, excess capacity to recover from failure.

(9) Options for introducing additional capacity to recover from a failure.

(10) Time of failure - this is not a primary variable but one where impact is reflected in all those above which are functions of time.

To be valid, a service availability measure must be sensitive to variations in these system parameters.

Based on these insights, the following observations emerge with respect to existing service availability measures:

(1) Existing measures reflect equipment performance and neglect the secondary effect of average queue dynamics.

(2) Because of this, existing analytic measures are not useful predictors of delay performance.

(3) Irrespective of (2), however, existing measures can be useful in monitoring roles, in a qualitative sense, because the system parameters which control the transfer function between system downtime characteristics and passenger delay characteristics can be assumed fixed.

(4) It would appear that analytic measures can be quantitatively useful within the context of a specific system, either in a predictive or monitoring role, if they were properly calibrated by a relationship between system downtime characteristics and passenger delay dynamics. Establishing such a calibration factor is not a straightforward process.

(5) Analytic models may be quantitatively accurate in a predictive or monitoring sense without the calibration referred to in (4) if the system capacity is inherently very large compared to the demands. (Station queues would be dissipated by the first vehicle following a failure, SRT $\approx$ 0). However, such a simplification should result from appropriate analyses and not be presumed a priori.

These conclusions are, however, secondary to a more significant conclusion. It is obvious that the values of the previously listed delay-influencing variables will be unique to a specific system in a specific application. It follows that an appropriate measure to direct the control of these variables will also be system and application specific. Hence, no single measure can be applied to different systems vying for a specific application or to a specific system vying for different applications. The parameters of the appropriate service availability measure may be similar (e.g., system and/or major subsystem failure rates and associated values for TTR); however, the allowable values for these parameters will be different.

As a result of this finding, the problem focus shifted away from attempting to find or define a universal measure to one of developing a methodology to derive the appropriate measure within the context of a given transport situation.

3.4 Development of a Methodology for Relating Passenger Delay Potential and System Failure Characteristics

As implied by the list of system variables which can influence the delays experienced by passengers, the delay mechanism is very complex. In a general sense, the evaluation of each of these variables on system performance requires some form of simulation; that is, some method of imposing failures on a normally operating system and "counting" the passenger delay impacts which result. By doing this, a sufficient number of times to effectively cover the range of system variables proposed, allowable values for these variables can be derived in response to a given passenger delay criterion. For complex systems, computer simulation techniques are mandatory. For "simple" systems, however, where normal operation can be "visualized", manual techniques can be used. These simple systems consist of shuttle loops, line-haul systems, connected loops, and similar types of transit systems. It is significant that the current operational AGT systems fit this category. Furthermore, it is likely that the near future system (the Downtown People Mover Systems) will fit this category.

This process and the resulting methodology are described in Appendix B. Briefly, the methodology involves the determination of passenger delay response to a "unit failure" which is a failure which completely stops the system for some arbitrary time period. This response effectively establishes a relationship among system downtime, passenger demands, origin-destination patterns, and excess capacity. Extrapolating this response to expected failure types, frequency, duration, and extent is accomplished by scaling procedures.

Because a given delay criterion can be met with a number of combinations of system parameters, the methodology developed in this program is appropriate as an evaluation tool: given a set of system failure character-

istics and other delay influencing parameters, the expected delay characteristics can be derived. It is within this context that the methodology is presented.

4. SERVICE AVAILABILITY CONTROL PROCESS

5. METHODOLOGY FOR RELATING SYSTEM FAILURE CHARACTERISTICS AND PASSENGER DELAY CRITERIA

6. AN APPLICATION EXAMPLE

As indicated previously, these three sections, as a set, constitute the guidelines for controlling AGT system service availability. When submitted as an interim report, this Task 3 and 4 report contained these sections. However, in the Final Report, these have been compiled separately as Volume III - Application Guideline Manual, and, hence, are not duplicated here.

APPENDIX A

ASSESSMENT OF EXISTING SERVICE AVAILABILITY MEASURES

As indicated in Section 3. , the definition of SAMS is most problematical when making the transition between off-normal performance and passenger delay parameters. This requires an understanding of the relationship between off-normal performance characteristics and their propensity to induce delays--in terms of frequency and duration. It is this single relationship which has received most of the attention in service availability discussions and activities. In Reference 1, numerous SAMs were identified. Extracting from this source, three types of measures were identified.

Type I. Measures of the classical availability form

$$\text{Availability} = \frac{\text{Successful time}}{\text{Total operating time}},$$

where the elements of the fraction may be expressed in terms of system hours, vehicle hours, or passenger hours.

Intended Meaning: Likelihood of being in a successful state of any random time during use.

Type II. Measures of the classical dependability form

$$\begin{aligned}\text{Dependability} &= \text{probability of success}\\ &= \text{availability x reliability}\\ &= \frac{\text{successful trips}}{\text{total trips}},\end{aligned}$$

where the elements are generally computed on a per trip basis for either vehicles or passengers.

Intended Meaning: Likelihood of not incurring a delay during a given period of use (generally one trip).

Type III. Measures of the expected delay form

$$\text{Expected delay} = \text{probability of delay x avg. duration of delay},$$

where the elements are generally computed on a passenger trip or vehicle-trip basis.

Intended Meaning: Average delay of a passenger on a typical trip.

As the term "availability" implies, Type I measures indicate the proportion of time in which a system (a subsystem) is operating in an unfailed state. Availability expressions are used as a proxy for the likelihood of a passenger not experiencing a delay in assessing the system or subsystem being analyzed. The classical form of the availability expression is

$$A = \frac{MTBF}{MTBF + MTTR}, \tag{A-1}$$

where A = Availability of system or subsystem being investigated,

MTBF = Mean time between failure - the average time, measured in terms of successful operating hours, between failure occurrence,

and MTTR = Mean time to restore, the average time to restore the system or subsystem to operating condition.

The most general use of such measures in assessing delay impact is to calculate availability at the system level; the results of which are assumed to represent the probability of a passenger not experiencing a delay at a station due to a system failure.

Type I measures do not indicate delay propensity once a passenger is using the system. These are rather determined by the reliability of the system,

$$R = e^{-\lambda T}, \tag{A-2}$$

where R = System reliability - probability that the system will remain in an operational state for a time T,

T = Time over which R is calculated,

and λ = Failure rate - the reciprocal of MTBF.

By establishing T as the trip time for a typical passenger and λ as the failure rate of the systems/subsystems utilized by this typical passenger, Equation (A-2) represents the probability of this typical passenger not experiencing a failure (and a resulting delay) en route.

Type II measures combine the station delay and en route delay characteristics into a single measure called trip dependability. The assumption used in this combination is that a passenger will not experience two failures during any given trip; hence, he will be delayed by either at the station or en route, but not both. Under this assumption, the classical dependability measure is

$$D = A \cdot R \tag{A-3}$$

where D = Trip dependability,

A = System availability as defined in Equation (A-1),

and R = System reliability as defined in Equation (A-2).

Types I and II measures address probability of success; that is, probability of experiencing no failure. The probabilities of delay can be developed by subtracting the above measures from unity.

$$P_{d/s} = 1 - A \, , \tag{A-4}$$

where $P_{d/s}$ = Probability of delay at station

and A = System availability as defined in Equation (A-1).

$$P_{d/er} = 1 - R \tag{A-5}$$

where $P_{d/er}$ = System reliability as defined in Equation (A-2).

$$P_d = 1 - D \, , \tag{A-6}$$

where P_d = Probability of delay at station or en route

and D = Trip dependability as defined in Equation (A-3).

Type III measures are tantamount to "risk" measures utilized frequently in safety assessment activities in that they comprehend not only the likelihood of experiencing a delay but also the consequences of a delay, if encountered, in terms of its duration.

As discussed in Reference 1, many formulations exist within each of these type categories. Each specific formulation reflects some particular characteristic of the system being analyzed or, perhaps, varying degrees of detail in the analysis. Regardless of the specific form of these expressions, they all have a common intent: to be a representative "model" relating system off-normal performance characteristics to passenger delay parameters. Initially, it was the intent of this project to define a list of criteria by which the various measures and models could be evaluated. Reference 2 contains this list. Ten criteria were offered as determinants of a "good" service availability measure.

(1) It should reflect passenger perception of service.

(2) It should reflect the performance of elements of a transit system over which the operator can exert control.

(3) It should be capable of measuring improvement in performance due to operator control action.

(4) It should lead to clear, unambiguous performance specifications which can be directly treated in the design and manufacture of a transit system. By the same token, it must be measurable with reasonable data collection efforts, free of ambiguity, during tests and revenue service.

(5) It should reflect system performance goals rather than subsystem requirements so as not to unnecessarily constrain the suppliers.

(6) Like (2) above, it should relate to the elements over which design control can be exerted (nominally, the failure characteristics of and restore strategies of the technology being proposed and/or constructed).

(7) It should be sensitive to small changes in design and/or operational parameters.

(8) It should have technical validity.

(9) It should be independent of any specific technology. If specifications are the result of allocating system level goals, this allocation should <u>not</u> favor one technology over another.

(10) It <u>must</u> be an effective communication tool among all elements of the transit industry.

It was anticipated that these criteria would be developed into independent, scalar measures against which all conceivable SAMs could be evaluated--to judge their applicability and utility as an effective service availability control parameter. It became obvious during the initial program efforts, however, that attempting to do so would serve no useful purpose. The main reason for this conclusion was that all of the above criteria address , implicitly or explicitly, two major underlying concerns.

(1) Do these "models" really represent the relationship between passenger delay parameters and system off-normal performance?

(2) Do these models and measures lend themselves to reasonable techniques for analytic determination and/or measurement?

As might be inferred, these concerns are basically incompatible. As models for relating passenger delay parameters and system off-normal performance become more faithful in replicating or predicting actual performance, they also become more complex and problematical in use. However, the second concern is met only if the first concern is adequately resolved. Hence, in the following subsection, attention is focused on the first concern above. The approach taken was to examine passenger delays accruing from simple failures in simple system situations to develop a basic qualitative understanding of the relationships between system off-normal performance and passenger delays. This understanding was then applied to existing measures to judge their ability to replicate this relationship. This approach was selected under the assumption that one or more measures would pass the test. This assumption proved to be incorrect. Hence, the passenger delay analysis took on an added dimension of defining the problem to scope remedial investigations. Section A.1 discusses, in qualitative terms, the delay dynamics of passengers to identify the system variables involved and the extent of the influence exerted by these variables. Section A.2 briefly "evaluates" existing measures. Section A.3 summarily discusses the problem of relating system off-normal performance and passenger delays.

A.1 The Dynamics of Failure-Induced Passenger Delays

Passengers of a transit system can experience delays induced by system failures in one of two ways.

(1) Directly, by being aboard a vehicle which is stopped or slowed by a failure in the system

(2) Indirectly, by being denied normal access to the system during a failure.

To examine these effects, a simple transit system is assumed. This system consists of a guideway loop serving four stations. Four example transport situations are assumed. The first two illustrate the delay characteristics for fixed trip demands originating at only one station as

affected by different failure types (full stop, reduced speed, and vehicle-out-of-service failure) under varying levels of system capacity. The third example illustrates the effects of adding demands at a second station with a full-stop failure. The last example extends the previous one to a situation where a full-stop failure affects only part of the system. As indicated previously, these examples are presented only to illustrate the qualitative influence of these variables.

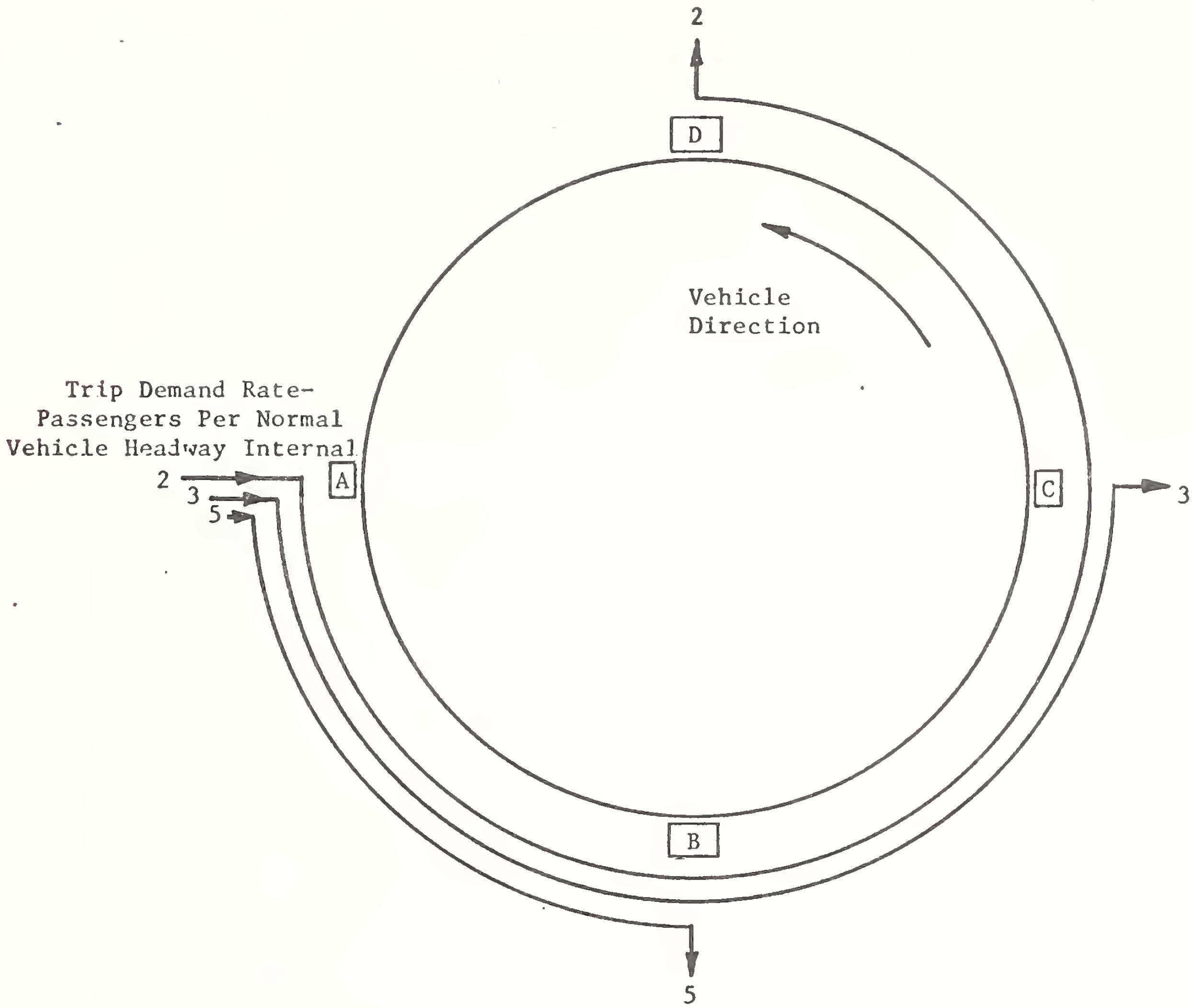
A.1.1 Example 1--Single Access Point, Full-Stop Failure

The transport situation for this example is illustrated in Figure A-1. This system consists of 4 stations, with Station A being the only station at which passengers board the system. The numerics in this figure represent the number of passengers entering a single vehicle at Station A and exiting at Stations B, C, or D. Vehicle velocity and station spacings are such that the vehicles appear to "index" between stations with station stops occurring at the same time, i.e., normal vehicle headway time equals transit time between station pairs.

This type of normal service is depicted by the train graph of Figure A-2. Under normal operation, passengers arrive at Station A in some random fashion (assumed to be a constant rate in this example) and are transported in accord with a fixed schedule. This relationship is depicted in Figure A-3 as the "demand" curve and the "normal supply" curve. In the normal situation depicted, a vehicle with a capacity of at least 10 passengers is required.

A failure is imposed on this system with a duration of 4 headway intervals.* This failure results in all vehicles stopping. During the down-time period, passengers continue to arrive at Station A, building up a queue of delayed passengers. After the failure is removed, this queue is dissipated at a rate dependent on the amount of extra passengers which can be carried per vehicle headway--the excess capacity in the system. This supply curve is denoted as the "off-normal supply" curve in Figure A-3 and was

* Generally, through this report, time units are expressed in terms of "normal vehicle headway intervals" or "headways". This normalizes time to those periods during which passengers expect to be served.



- Flow Rates Based on a Per-Vehicle or Per-Headway Basis
- System Utilizes 4 Vehicles, Equally Spaced in Time
- Passenger Demand is Constant for Duration of Analysis Period

FIGURE A-1. EXAMPLE SYSTEM FOR ESTIMATING DELAY PARAMETERS

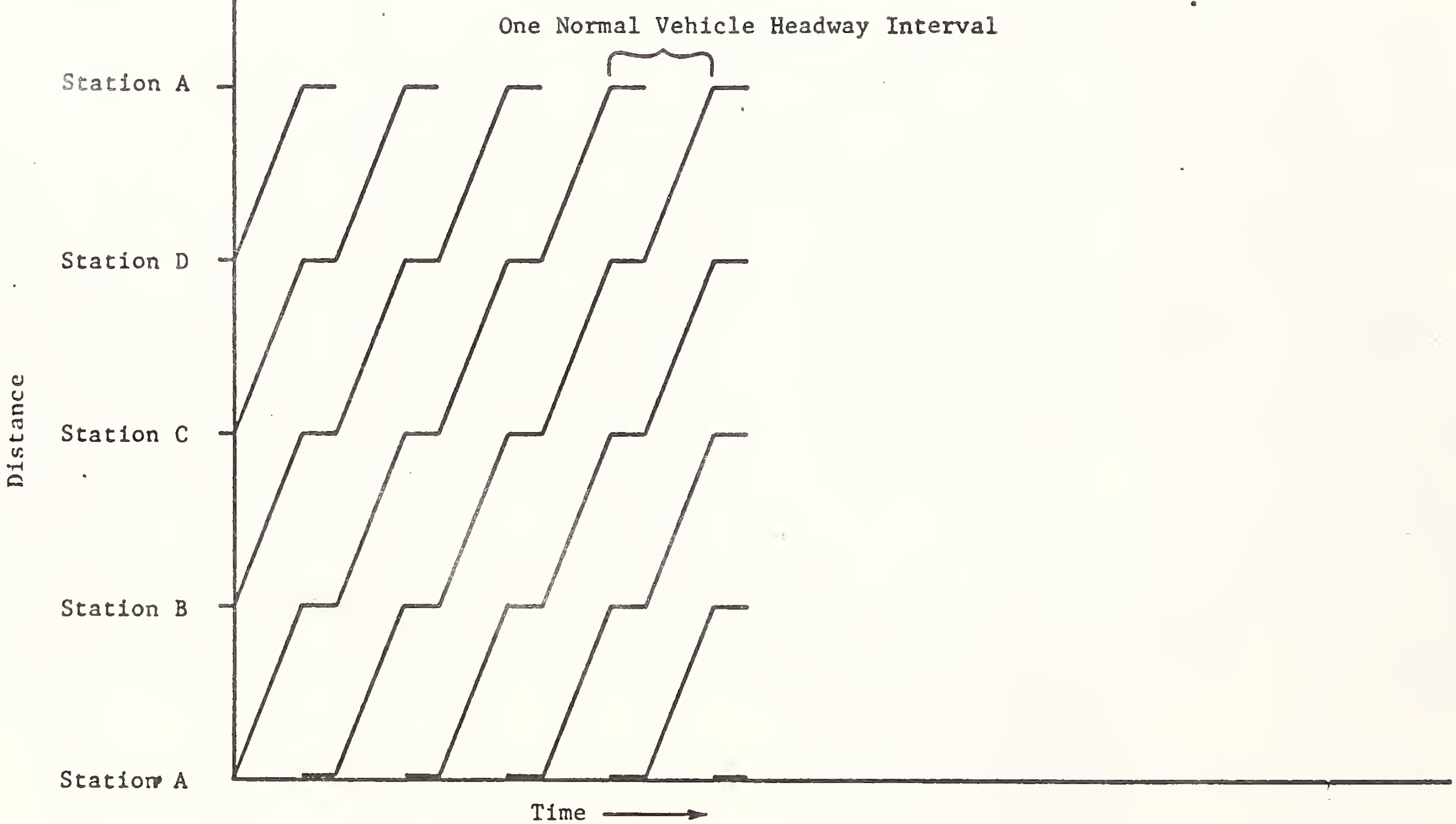


FIGURE A-2. TRAIN GRAPH SEGMENT ILLUSTRATING NORMAL SPEED/DISTANCE PROFILE OF VEHICLES FOR EXAMPLE SYSTEM

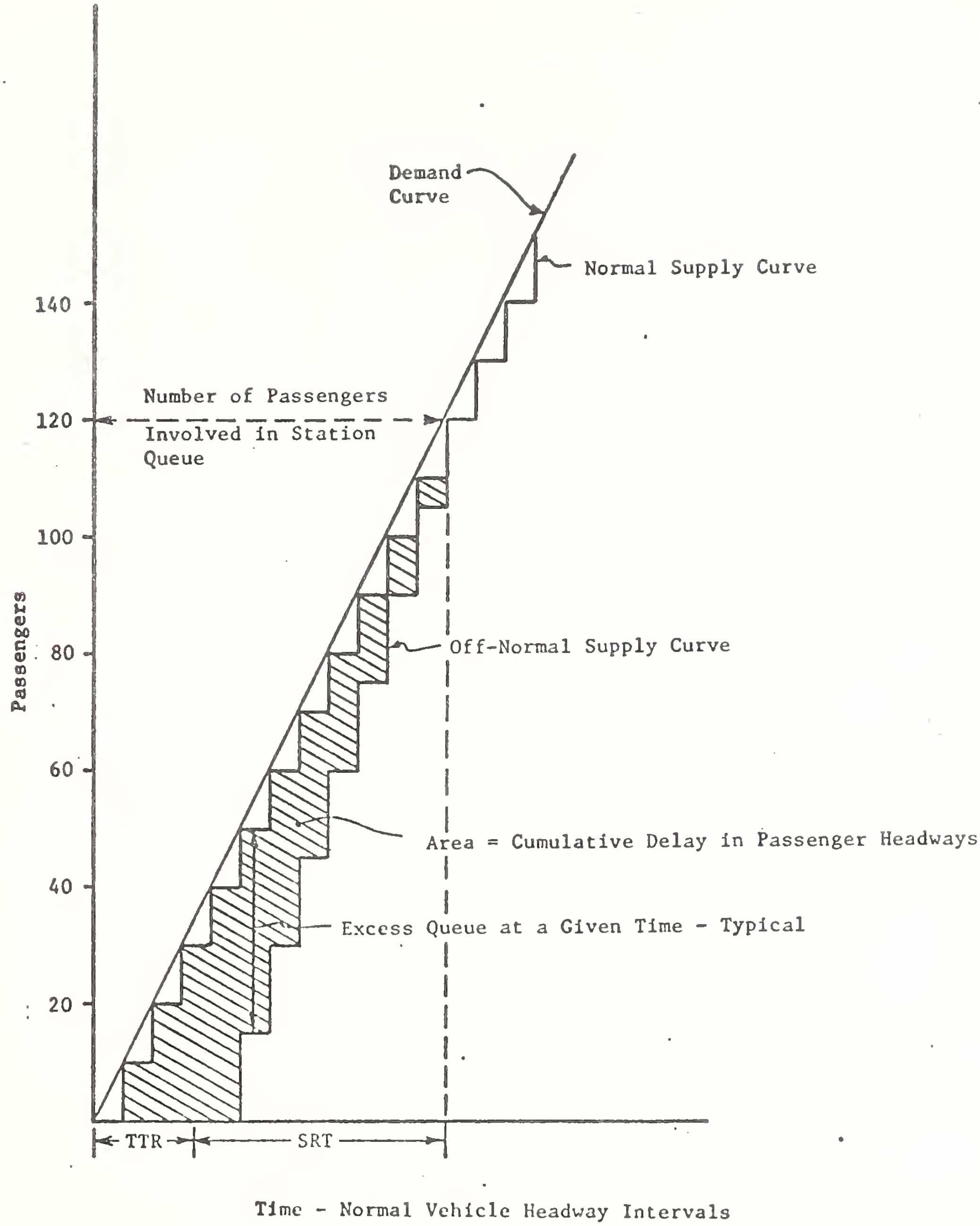


FIGURE A-3. DELAY IMPACT OF FULL-STOP FAILURE

generated under an assumption of 15 passenger vehicles or a 50 percent excess capacity* in the system.

Several definitions are indicated in Figure A-3 which will be used in subsequent analyses.

(1) Cumulative passenger delay is represented by the area between the normal supply curve and the off-normal supply curve. This representation does not include the normal waiting time of a passenger under normal service conditions. This is desirable since the issue is failure impact.

(2) A "time to restore" (TTR) is indicated which represents the perceived system downtime at that station. This term stands for the time required to make the equipment aspects of the system operating normally, e.g., vehicles following normal speed/distance/headway profiles. This is equivalent to "mean time to restore" (MTTR) in most availability expressions.

(3) A "service restore time" (SRT) term is indicated which represents the additional time (beyond TTR) required to reduce station queues and, hence, their effects to normal conditions.

(4) The number of passengers involved in the delay event is indicated. It is not obvious exactly how many of these passengers were actually delayed. This depends on how the passengers orient themselves within the station and how aggressive they are in boarding available vehicles. It is assumed here that the queue discipline would be a first-in-first-out process, which results in everyone experiencing some delay, with the exception of a few near the end of SRT.

* As used in this report, excess capacity is referenced to a particular station. It is equal to the number of passengers per unit time which the system is capable of transporting from that station compared to the normal demand generated at that station. The unit of time is one normal vehicle headway interval. Hence, if vehicles are traveling at normal speeds, excess capacity is identical to the vehicle capacity minus the number of passengers normally on board as a vehicle leaves the station being referenced.

Figure A-3 illustrates the delay parameters for passengers delayed at the station. In addition, passengers are delayed en route. Under the failure characteristics assumed, the number of passengers experiencing an en route delay would be equal to the per headway link loadings. Specifically, 10 passengers would be delayed between Stations A and B; 5 would be delayed between Stations B and C; and 3 would be delayed between Stations C and D. Neither the number of passengers delayed nor the duration of delay of en route passengers is affected by SRT; i.e., en route delays depend strictly on failure frequency and TTR.

A.1.2 Example 2--Single Access Point, Slow-Speed/ Vehicle-Out-of-Service Failure, Varying Excess Capacity

To illustrate the effect on delay parameters of different failure types, for the example system of Figure A-1, Figures A-4 and A-5 are presented. Figure A-4 shows the effect of a failure which results in all vehicles running at half speed. Two values of excess capacity are used: 50 percent and 100 percent. Figure A-5 shows the effects of a failure which results in a vehicle being temporarily removed from service. Again, two values of excess capacity are illustrated. In each of these failure situations, the higher values of excess capacity result in repeated, short-term delay events. These result from the lack of schedule adherence when experiencing these failures. When the lower values of excess capacity prevail, however, a more serious problem is encountered. Not only are the short-duration delays in existence but there is also a loss of average capacity to the point where the average supply capability of the system is insufficient to meet the accumulated backlog of transport-demand--irrespective of the short-term repetitive effects. This capacity loss induces the most serious effects on passenger delays.

Like Figure A-3, Figures A-4 and A-5 address only the platform delay parameters induced by the failures examined. For the slow-down failure, all passengers on the guideway at the time of the failure will experience a delay en route to their destination, the duration of which is a function of the passenger's location relative to his destination at the time of the failure. Additionally, all passengers departing Station A during TTR will

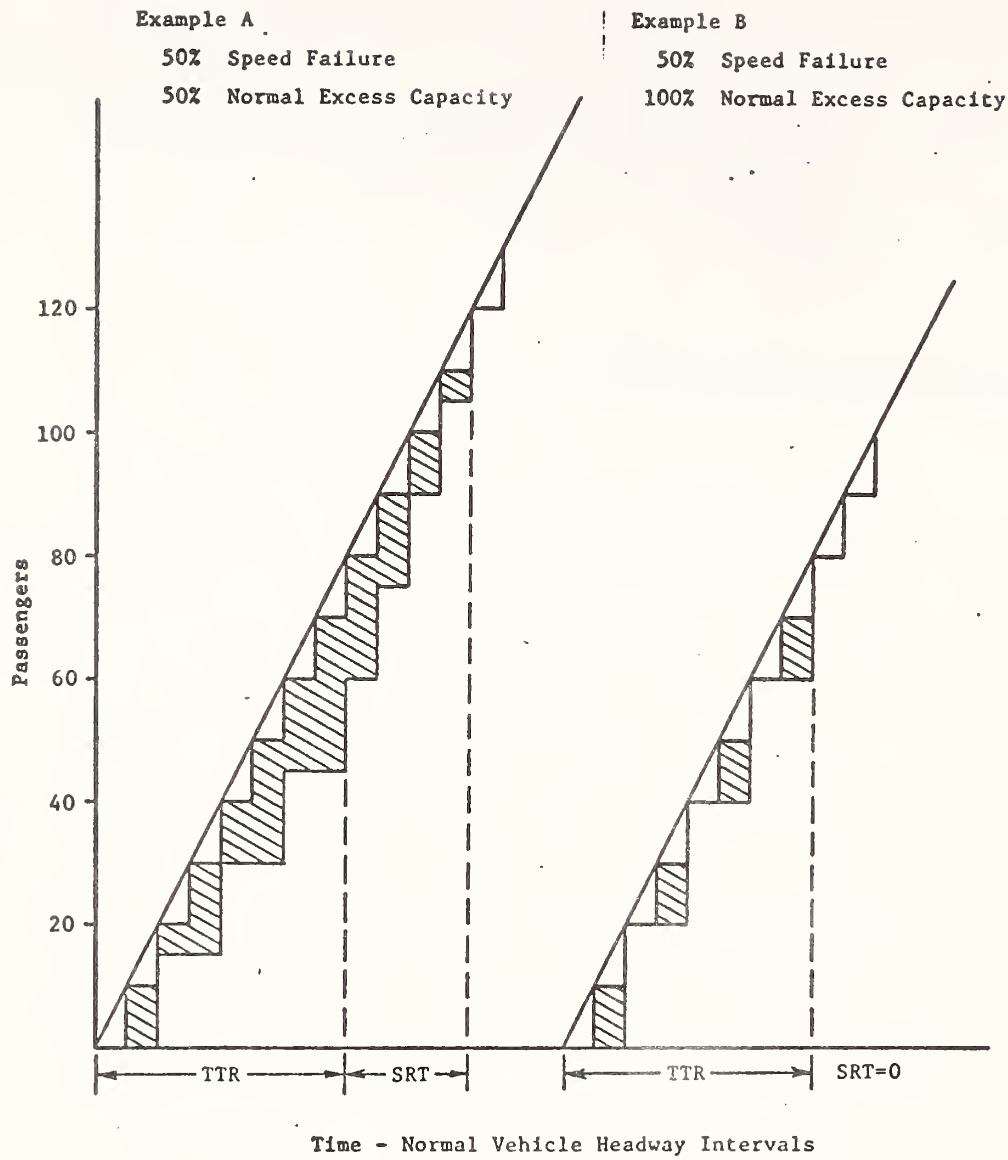


FIGURE A-4. DELAY IMPACTS OF SLOW-SPEED FAILURES

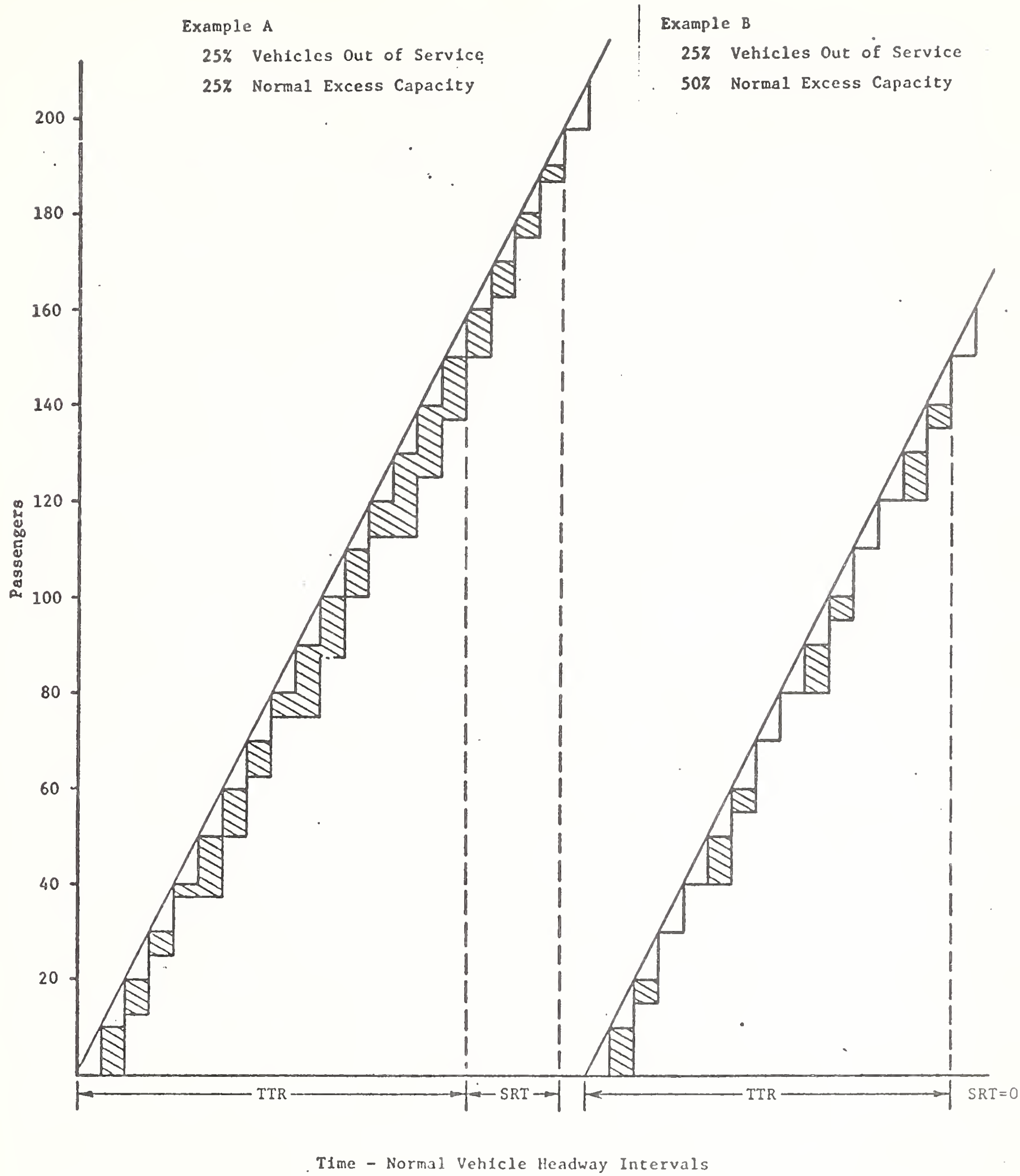


FIGURE A-5. DELAY IMPACTS OF VEHICLE-OUT-OF-SERVICE FAILURES

experience a delay, the value of which is a function of trip length and departure time relative to the end of TTR.

For the vehicle-out-of-service failure, assuming the vehicle can be moved to a station or siding for removal or repair without interfering with other vehicle movements, the only en route delays are those experienced by the passengers aboard the disabled vehicle.

Even for these simple examples, the complicated nature of the relationship between passenger delay parameters and system failure characteristics is emerging. General delay parameters are neither predictable nor measurable by simple relationships of failure rates and restore times. System failure rate is obviously a prime factor in assessing performance over some period of time in that this characteristic defines the number of failure events to be expected. The effects of each failure event, however, are significantly influenced by the impact of the failure on system operation (e.g., full stop, slow speed, etc.) and the relationship between the transport demands and the available system capacity--specifically the excess capacity.

A.1.3 Example 3--Two Access Points, Full-Stop Failure

This complexity is further illustrated when the simple example of Figure A-1 is expanded by incorporating passenger demands at another station. Figure A-6 illustrates this transition. As shown, passenger demands are imposed on Station C. As before, all numerics are demands per normal vehicle headway interval. All other assumptions are identical with those in the previous example with the exception of vehicle size. For the Figure A-6 example, maximum link loading of 14 passengers per vehicle headway occurs between Stations A and B. Thus, under normal operation, a 14-passenger vehicle would be required. To maintain some similarities to the previous example, a vehicle size of 19 passengers is assumed. This results in a normal excess capacity of 50 percent relative to Station A demands and 112 percent relative to Station C demands.

Figure A-7 illustrates the effects of a full-stop failure with a duration of 4 headways on resulting passenger delay parameters. The Station A curves can be compared directly to Figure A-3. As can be seen, the delay performance is considerably altered. While, under normal conditions, a 50 percent

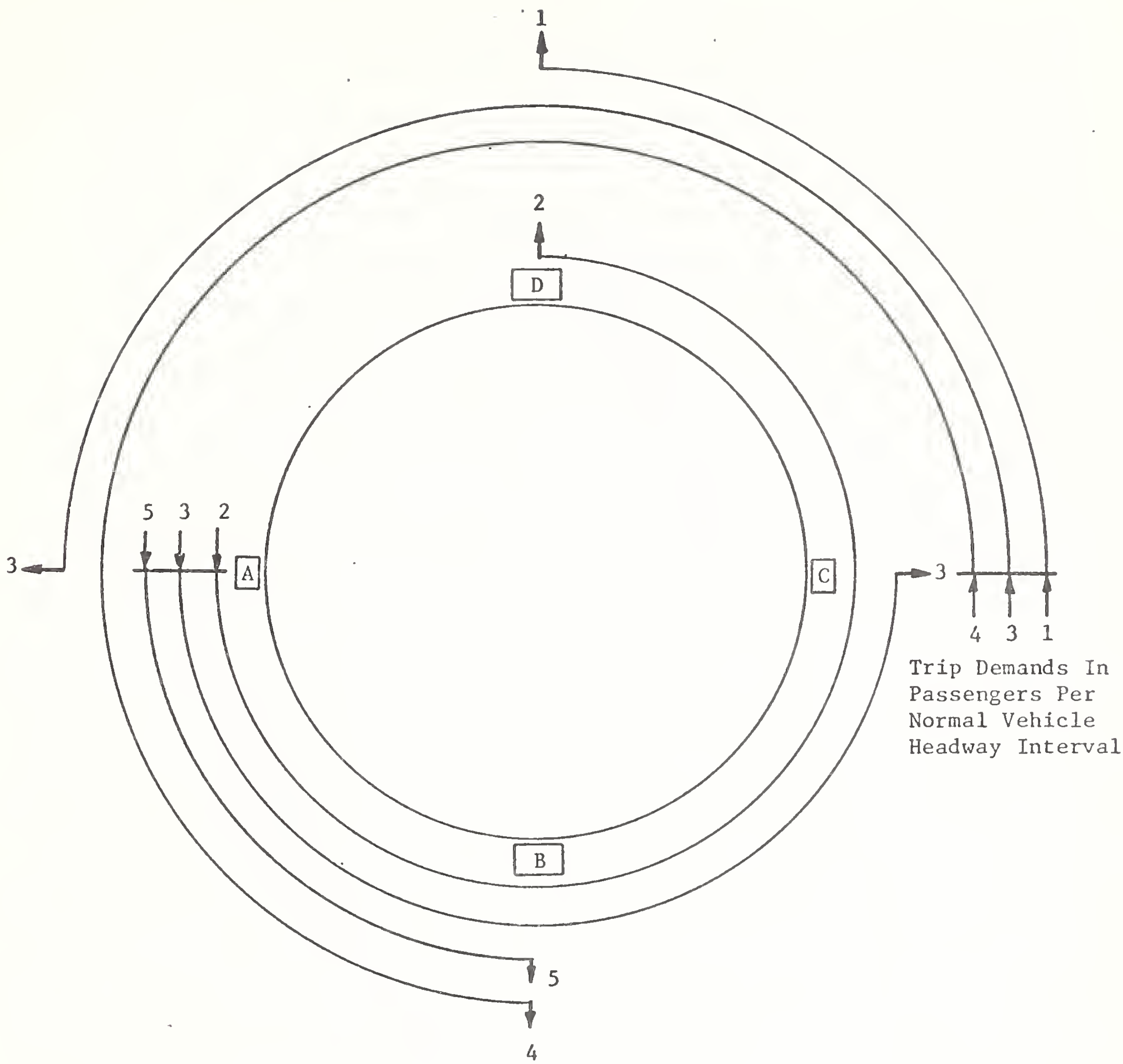


FIGURE A-6. EXAMPLE TRANSIT SYSTEM TO ILLUSTRATE EFFECTS OF MULTIPLE STATION INTERFERENCE ON QUEUE DISSIPATION PROCESSES

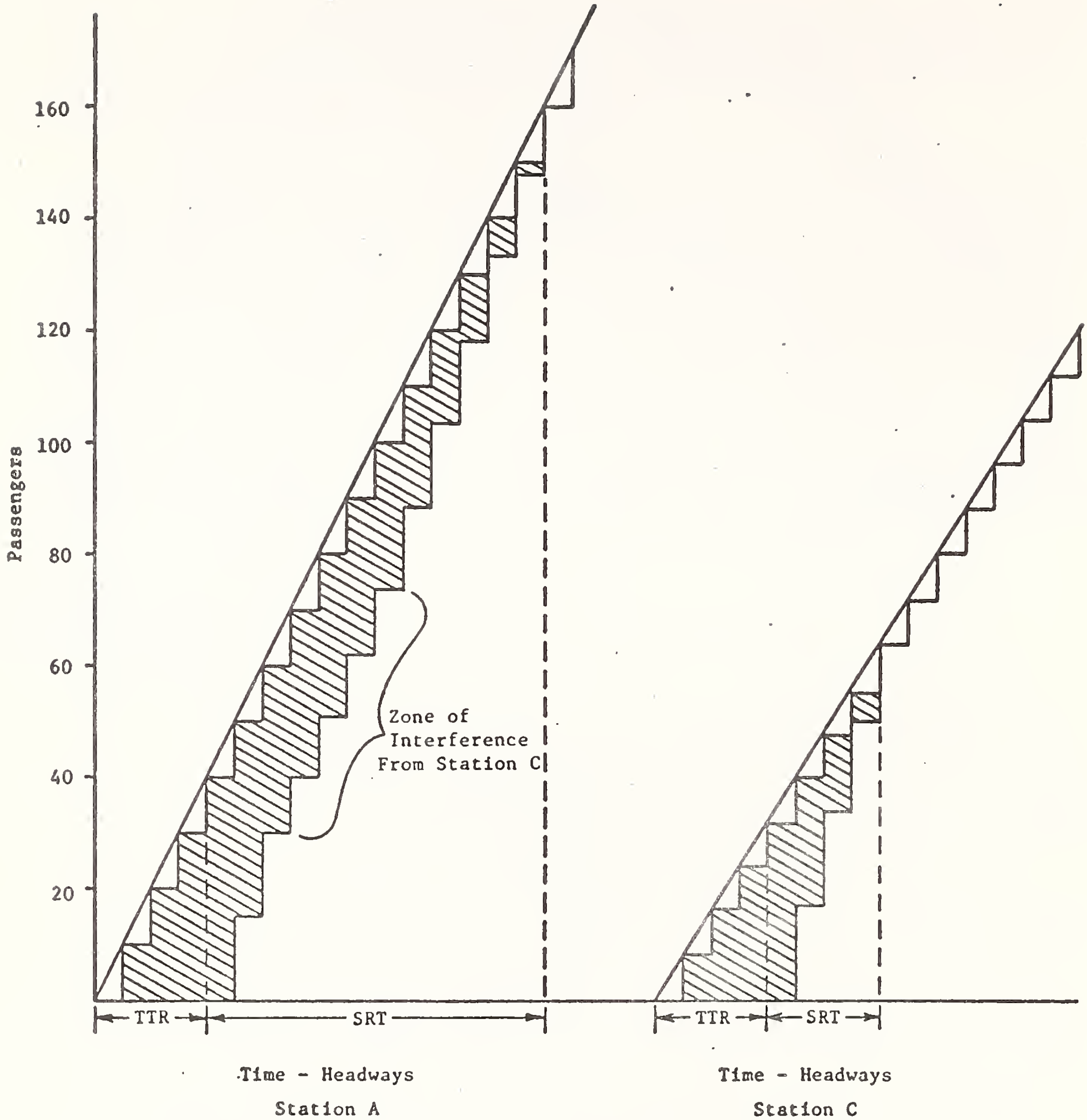


FIGURE A-7. EXAMPLE OF QUEUE DISSIPATION INTERFERENCE AMONG STATIONS - FULL-STOP FAILURES

excess capacity is available at Station A, the usable portion of this following a failure is reduced because of competition from Station C--specifically the increased trip volume through Station A as the delayed queues from C are dissipated. (This interference zone is indicated in Figure A-7.) Hence, for any given station, the addition of demands elsewhere in the system alter not only the SRT value but also the shape of the off-normal supply curve within SRT. Without resorting to further examples, the degree of alteration is a function of the trip patterns (i.e., origin-destination patterns) generated at these influencing stations.

A.1.4 Example 4--Two Access Points, Full-Stop Failure, Partial System Affect

As a final example in this section, the effects of a failure which renders only part of the system unusable are examined. In this example, it is assumed that the stations can be bypassed and a failure is experienced at Station B such that access is denied. To emphasize the effects, it is assumed that this failure exists for 8 headways. It is assumed that station announcements at Stations A and C would indicate the closure of Station B and would further request all passengers at that station to wait on the platform until service has been restored--and that passengers abide by this request. As a result, during TTR, queues build up at Stations A and C composed of passengers destined to Station B. After Station B is opened, these queues begin dissipating with a rate dependent, as before, on the available excess capacity of the system. Because of the dependence of excess capacity on trip patterns, a queue dissipation law has to be assumed. For this example, it is assumed that preference is given to passengers in the delay queue. Under this assumed service law, the off-normal supply curves indicated in Figure A-8 result.

In this example, as well as those depicted in Figures A-4 and A-5, no attempt was made to determine the number of passengers delayed. For the full-stop failure (Figures A-3 and A-7), the number of passengers delayed can be approximated by the product of demand rate* of the quantity (TTR + SRT). As

* Demand rate is used in this report to indicate passenger trip demands at a given station per unit time. The appropriate unit of time is generally one normal vehicle headway interval. It may vary from this in certain expressions to maintain dimensional consistency.

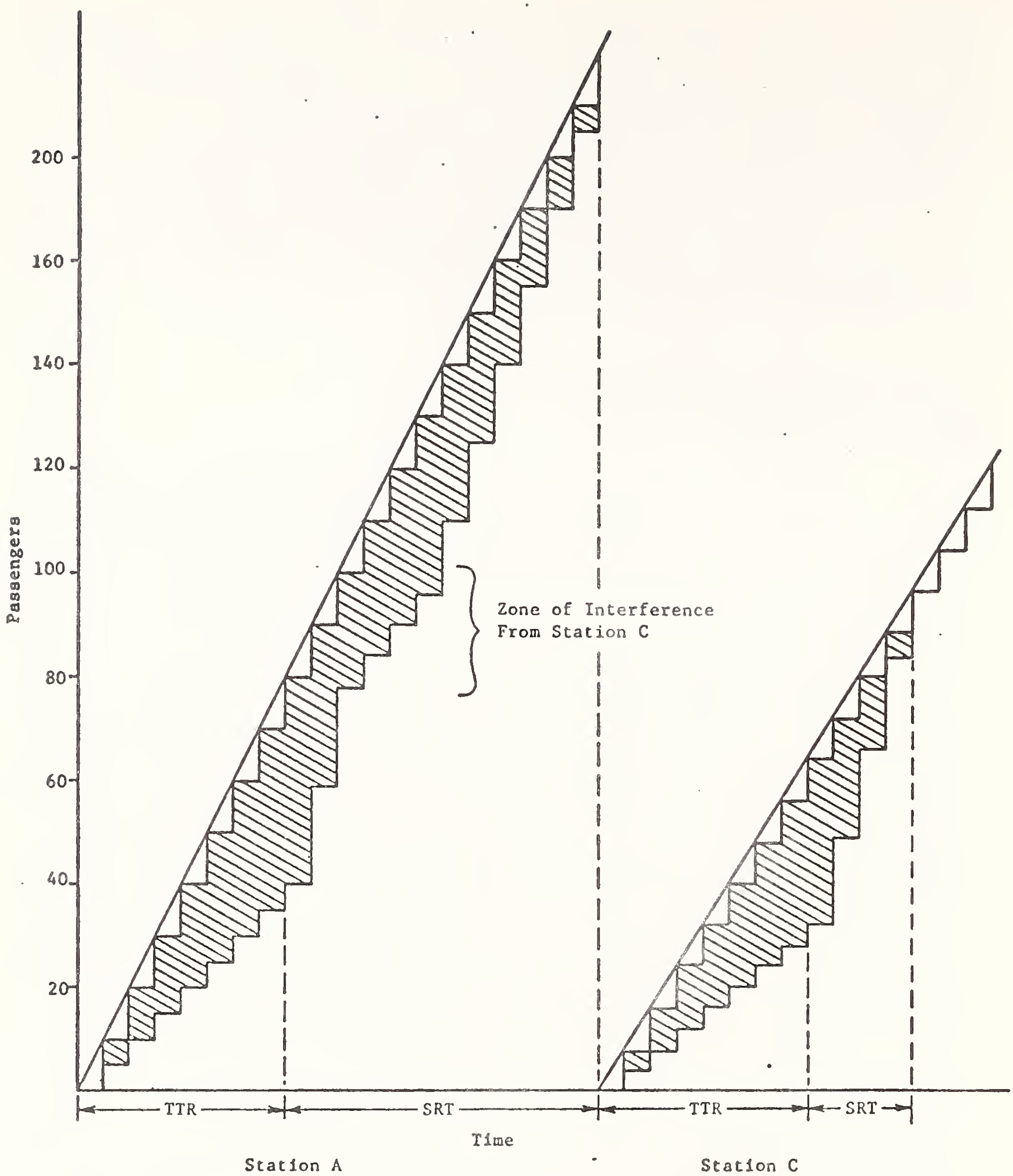


FIGURE A-8. ILLUSTATION OF QUEUE DISSIPATION INTERFERENCE AMONG STATIONS - PARTIAL SYSTEM FAILURES - STATION B UNAVAILABLE

was noted in the discussion of Figure A-3, the error associated with this approximation occurs near the end of SRT, where passengers entering the station can board a vehicle at their normal time, albeit a vehicle more crowded than normal. When failures occur which result in degraded service, the type of error becomes much more significant and its determination can become a complicated probabalistic exercise.

In summary, it can be said that the mechanism of failure-induced passenger delay is a complex process. The following points are specifically stressed:

(1) The influence of a system failure on passenger delays can exist long after the transit system equipment is restored to normal operating levels--an SRT exists.

(2) The SRT value is influenced by the inherent excess capacity in the system.

(4) The ability of this excess capacity to minimize delays is affected by the demands, in terms of quantity and trip profile, generated to upstream stations.

(4) Delay effects are related to the type of failure and its effects on system operation.

A.2 Assessment of Existing Measures

Earlier in this Appendix, general types of performance measures were introduced. These measures, as indicated by their form, consist of system off-normal performance parameters, combined in such a way as to compute a value which is interpreted in terms of some passenger delay parameter. In concept, therefore, they fulfill one requirement of SAM 2, as defined in Section 3.2. As pointed out earlier, the primary question regarding these measures is the validity of the interpretation.

Each individual measure discussed in the literature, for practical reasons, cannot be specifically analyzed to check the validity--nor is this necessary. All such measures can be interpreted as vehicle motion oriented--generally relating to schedule adherence capabilities. For example, Type I measures, while generally assumed to define probability of no delay at the station, actually define the probability that the scheduled service at a

station will be maintained. Type II measures reflect probability with the added probability that the vehicle will not experience a delay en route. Equating these measures to passenger delays assumes a one-to-one correspondence between passenger "schedule adherence" and vehicle schedule adherence. This assumption is obviously valid for en route delays. As pointed out in Section A-1, however, it is generally invalid for station delays. Normal service at a station may not be restored for some time--SRT--after the system equipment is operating normally. The only instance of 1:1 correspondence is when the excess capacity of the system is sufficiently large to clear the delayed passenger queue in one normal vehicle headway interval.

Figure A-9 is a reproduction of Figure A-3 with an area highlighted to illustrate the passenger pattern which is reflected in classical Type I service availability measures. Neglecting the step differences which exist at the lift portion of this pattern (which would get smaller as the headway/TTR ratio gets smaller), it can be seen that classical measures treat only the equipment downtime patterns and deduce passenger delay parameters by direct correspondence to this equipment downtime. Secondary delays due to queue dissipation effects are not considered.

It is important to note that fault of existing measures--assuming they represent off-normal performance--is their lack of quantitative correspondence. In any given situation, <u>any</u> measure which accurately defines system off-normal performance has to reflect passenger delay propensity in a qualitative sense; i.e., improving off-normal performance will improve passenger service. Hence, for an existing system, existing measures are useful in monitoring changes in performance, even though the actual values measured cannot be directly related, in a quantitative sense, with passenger delays.

Based on the above arguments and examples, the following observations emerge with respect to existing measures:

(1) Existing measures reflect equipment performance and neglect the secondary affect of average queue dynamics.

(2) Queue dynamic effects are controlled both during the buildup phase and the dissipation phase by the relationship of system capacity and passenger demand, in terms of both trip generation demand and origin-destination patterns.

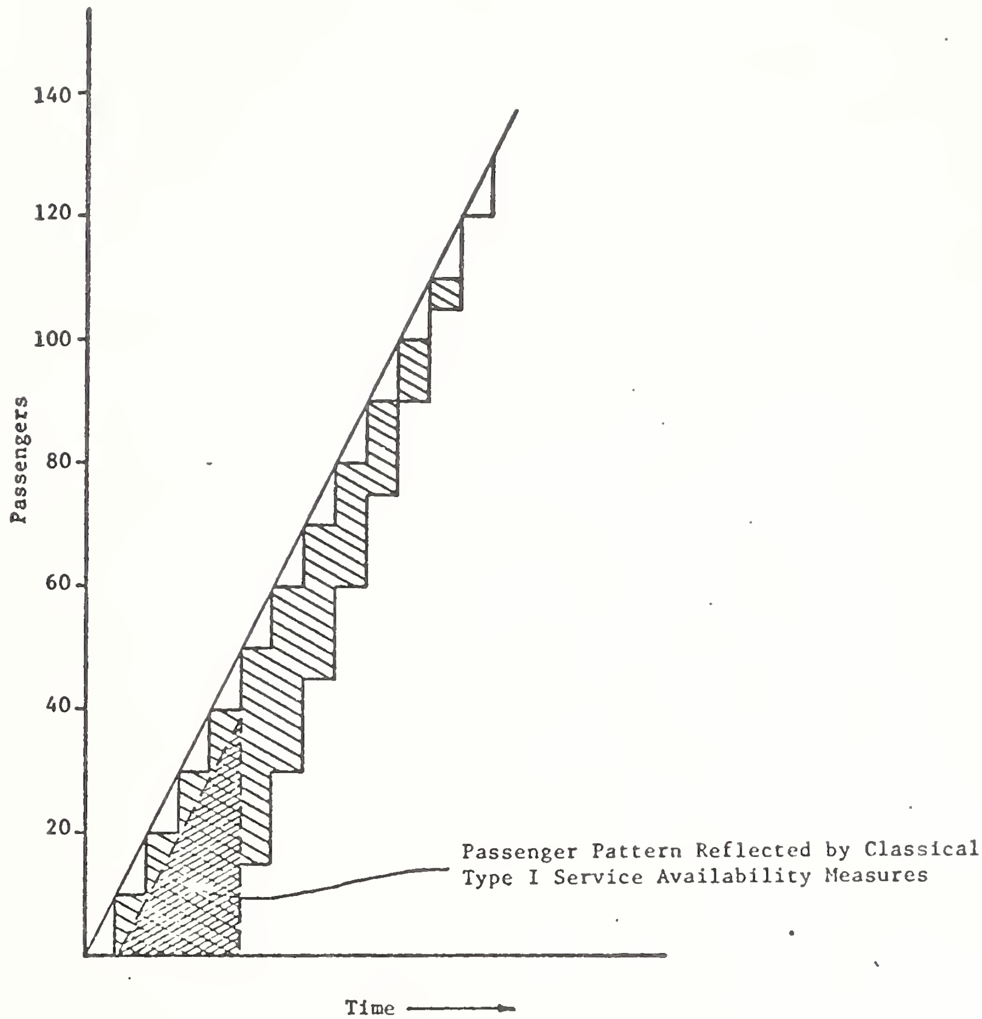


FIGURE A-9. ILLUSTRATION OF RELATIONSHIP OF DELAY PARAMETERS PREDICTED VIA CLASSICAL AVAILABILITY MEASURES WITH ACTUAL DELAY PARAMETERS

(3) Because of (1) and (2) above, existing analytic measures are not useful predictors of delay performance, regardless of the scope of the system.

(4) Irrespective of (3), however, existing measures can be useful in monitoring roles, in a qualitative sense, because the system parameters which control the transfer function between system downtime characteristics and passenger delay characteristics can be assumed fixed.

(5) It would appear that analytic measures can be quantitatively useful within the context of a specific system, either in a predictive or monitoring role, if they were properly calibrated by a relationship between system downtime characteristics and passenger delay dynamics. Because of (2) above, establishing such a calibration factor is not a straightforward process. Computer simulation techniques would probably be required to effectively manipulate the variables involved.

(6) Analytic models may be quantitatively accurate in a predictive or monitoring sense without the calibration referred to in (5) if the system capacity is inherently very large compared to the demands. However, such a simplification should result from appropriate analyses and not be presumed a priori.

A.3 General Conclusions

The previous section results in the conclusion that no existing measure is sufficient to act as a control agent because none reflect the impact of SRT on delay dynamics. It would seem obvious that calibrating these off-normal performance measures to reflect SRT would be appropriate. This is correct. However, as will be discussed in Appendix B, this is not a simple process. Techniques are developed to accomplish this analytically for relatively simple systems (e.g., loops, shuttles, line-haul systems, etc.) in which "normal" service can be visualized and analyzed. For systems

which are characterized by random routes, demand response, and other random features such that computer simulation is required to estimate "normal" performance, the relationship between off-normal performance and passenger delays will similarly require computer simulation.

APPENDIX B

METHODS FOR ESTIMATING FAILURE-INDUCED DELAY PARAMETERS

The deficiencies in the ability of equipment failure performance to act as a reasonable proxy for induced passenger delay effects led to the initiation of a limited parametric analysis aimed at the full definition of the relationships between system failure characteristics and the resulting passenger delay effects. As pointed out in Appendix A, the major problem is one of properly assessing the impact of system failures on queue buildup and dissipation in the stations, the effect of which is a denial of normal access to the transit system. Appendix A illustrated the ability of equipment failure characteristics to define the occurrence of delay events but not the full effect of such events as experienced by the passengers. As pointed out in Appendix A, reliability characteristics of transit systems are sufficient to define the propensity of the system to induce delays during a given trip. The delays induced at stations are, however, only partially related to the system failure characteristics. Identified, but not quantified, were certain parameters which appear to control this relationship, such as

The passenger trip demands at individual stations

The excess passenger carrying capacity available to dissipate queue buildup at stations

The type of failure

The interaction of trip demands among stations.

A significant conclusion of Appendix A was that while no existing availability measure is sufficient, in itself, to act as an effective passenger delay control measure in a quantitative manner, all existing measures can perform this function in a qualitative sense.

The implication of this conclusion is that existing measures, which accurately reflect system off-normal performance, can be augmented to account for the effects of SRT, the service restore time. As indicated in Appendix A and further supported in the following section, the difficulty with this approach is the system-specific nature of this augmentation. At this time, we cannot see a way to formulate a general-purpose "augmented" service availability measure. Rather, each specific system situation will require its own

measures, as appropriate, to its specific characteristics. Hence, the emphasis of the following subsection is placed on methodology concepts involved in relating equipment off-normal performance and passenger delay parameters.

This methodology was developed by generalizing and testing results of over 50 manual simulations of simple transit system situations which were examined completely or partially as needed to test various hypotheses. The use of the resulting methodology forces the analyst to understand the delay dynamics at work in his specific system situation; irrespective of the approximate nature of the values determined, this understanding is considered to be an important attribute of the methodology developed.

The remainder of this section discusses this methodology and its supporting concepts.

B.1 Methodology Concept

When one views the complex, dynamic nature of transit system operations, the variations and probabilities nature of failure performance, including type, frequency, duration, location, and time, and the probabilistic nature of passenger trip demands, the relationship between passenger delay potential and system failure characteristics at the system level cannot be viewed in any sense other than mathematically intractable. To quantify this relationship requires the subdivision of the system/passenger elements into units which are amenable to analysis. There are two basic approaches for accomplishing this. The first is to select a random, average, or typical passenger and "follow" him through the system, documenting his exposure to system failures and the impact of the exposure on his normal progression on an average or typical trip in terms of likelihood of encountering a delay event and the duration of this encounter, if it occurs. If the system performance criteria are stated in terms of allowable delays to the typical passenger, comparison can be made at this point. System-level performance can be deduced by extrapolation techniques for critical comparison at this level. This approach is inherently complex because of the need to define a "typical" passenger and the fact that his exposure to delay events is a function of location and time.

The second approach is to select locations and "watch" these over a period of time, accumulating data on passenger delays encountered at that location as a result of system failures. If these "observation ports" are properly selected to be the complete or typical set of locations where delays can accrue, system-level delay performance can be deduced. By comparing delay performance at these ports with passenger throughput at these ports, delay parameters for individual passengers can be deduced. This approach explicitly isolates location effects, is tolerant of using average values for failure characteristics and passenger dynamics, and eliminates the problems of identifying the "typical" passenger trip.

The methodology developed in this program utilizes this second approach. The observation ports are taken to be those points where passenger delays accrue--specifically stations and guideway links. Briefly, the methodology involves postulating failures and computing their impacts on delays at each of the analysis locations in terms of the number of passengers delayed and the duration of the delays experienced. These are then summed in accordance with the expected frequency of occurrence of the failures to arrive at system-level performance values for some selected period of time. This delay "data set" can then be manipulated, as necessary, to interface directly to the form in which delay criteria were established.

In the development of this methodology, it was recognized that passenger delay frequency and duration are affected by the following variables:

(1) Failure type - classified by the effect of the failure on the ability of the system to deliver required capacity in the vicinity of the failure. Three types are considered: (a) failures which result in a blockage, (b) failures which result in operations at velocities less than the normal velocity, and (c) failures which result in operations with less than the required number of vehicles.

(2) Failure rate - the expected number of failures in some unit of time.

(3) Failure duration - the time during which the failed state exists.

(4) Failure location - the location of the failure relative to the general system configuration. This is important where failure tolerance is provided.

(5) System failure tolerance - the ability of the system to limit the impact of certain failure situations by bypassing or otherwise disconnecting the failure affected area. This feature determines the extent to which a specific failure disturbs total system performance.

(6) Passenger trip demands - in terms of the quantity of trip requests per unit time.

(7) Trip origin-destination patterns.

(8) System capacity - more appropriately, excess capacity to recover from failure.

(9) Options for introducing additional capacity to recover from a failure.

(10) Time of failure - this is not a primary variable but one where impact is reflected in all those above which are functions of time.

In the development of the methodology, the effect of certain of the above variables were analyzed in isolation and then scaled to incorporate effects of other variables to arrive at system-level characteristics.

The following subsections discuss the specific methodology development procedures and results. Because of the different delay mechanisms at work for station delays and en-route delays, these are treated separately. This is also desirable because the delay criteria may treat these two delay situations differently. Also, in the case of en-route delays, a separation is made between general delays and stoppages--again to afford direct interfaces to possible exclusive delay criteria.

B.2 Station Delay Methodology

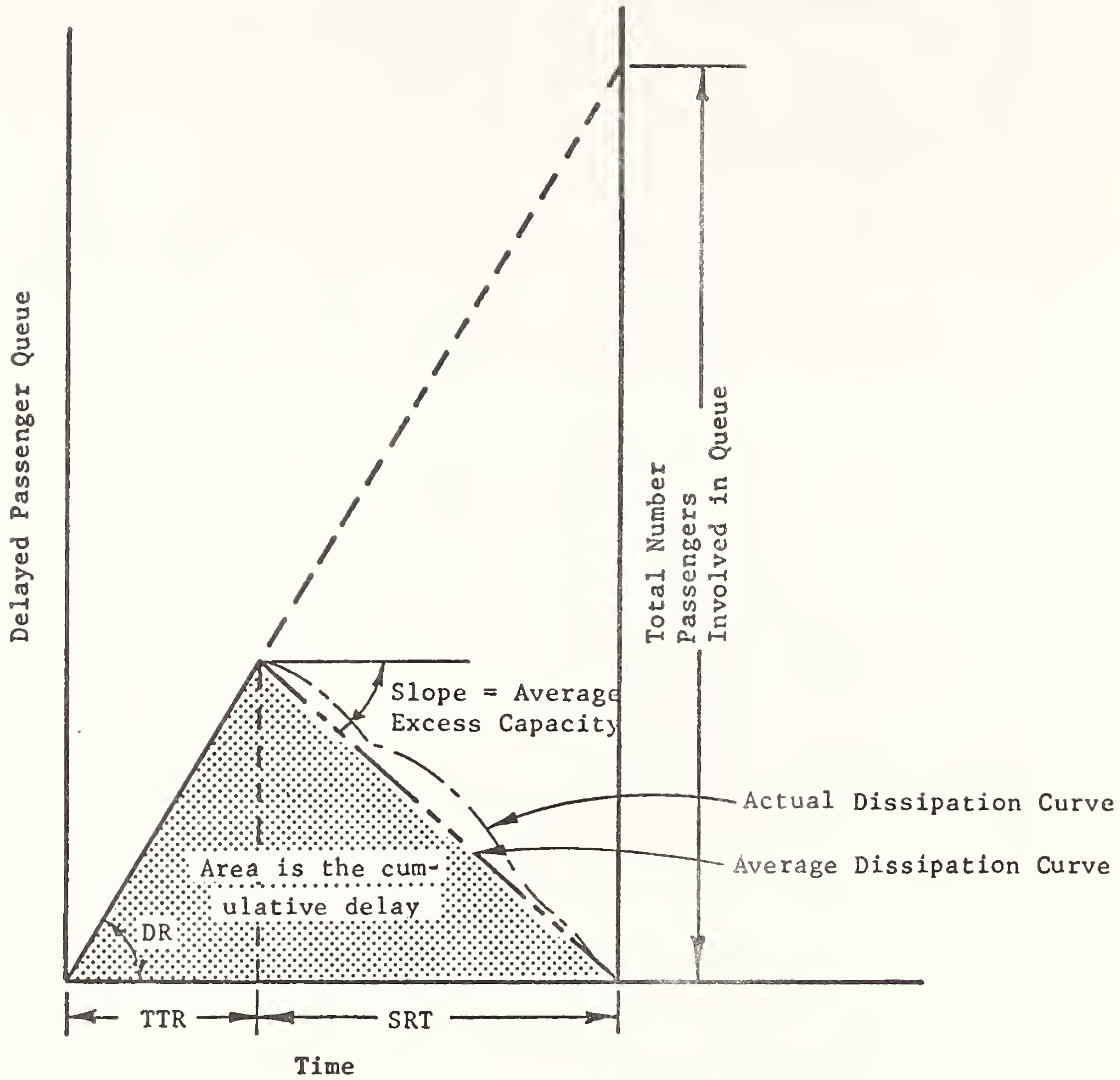
Under normal conditions, passengers enter a station with some expectation of a vehicle arrival time. If, because of a failure, vehicle service to that station is denied, these demands accumulate. As illustrated in the example of Figure A-1*, the magnitude of this accumulated demand is directly related to TTR, the time required to restore service to that station. Also, as illustrated in Figure A-3, after service is

* Figures with prefix "A" are included in Appendix A.

restored, this accumulated demand is reduced to zero at some rate which depends on the excess capacity of the vehicles which service that station. This delayed passenger queue can be graphically illustrated by figures such as Figure A-3. Because we are concerned only with the delay parameters, however, an alternate form is desirable, as illustrated in generalized form in Figure B-1. Such a representation can be used to define the number of passengers involved in the delay queue, as well as the cumulative delay incurred by these passengers--the area enclosed by the curve. In this figure, the queue dissipation curve is purposely indefinite: it represents the "unknown" quantities associated with passenger delays at stations. Specifically, these unknown quantities are the time to reduce the delayed queue to zero (SRT) and the shape of the delayed queue curve during the dissipation process. Defining these unknowns is the subject of the next subsection.

In the numerous examples analyzed in this program, it was found that, for a fixed value of passenger demand rate, system capacity, and passenger origin destination pattern, the delayed queue curve for a given station for varying values of TTR are geometrically similar. This implies that the delay parameters derived from system response to some arbitrary value of TTR (a "unit failure") can be considered to represent a "characteristic response" function; from which, delay parameters for other values of TTR can be drived via simple scaling laws. For example, in a given situation, the only independent variable identified in Figure B-1 is TTR. As TTR is varied, geometrically similar "triangles" result. Hence, the number of passengers delayed is proportional to TTR while the area enclosed by the "triangle"--the cumulative delay experienced by these passengers--is proportional to TTR^2.

It was found that such relationships exist for other failure types (e.g., slow speed), hence they form a basis for estimating the expected delay parameters for any specific system. The procedure would involve the determination of the characteristic response of the system and, using this as a base, generating the expected system response for actual expected values of TTR and failure rates. This approach is depicted in Figure B-2 as it might appear in the transit system design ohase. In an operational monitoring phase, the approach would be essentially the same excpet that observed values of TTR and failure rate would be used to assess performance.



DR = Passenger Service Demand Rate

FIGURE B-1. GENERALIZED REPRESENTATION OF PASSENGER DELAY PARAMETERS FOR STATION DELAYS

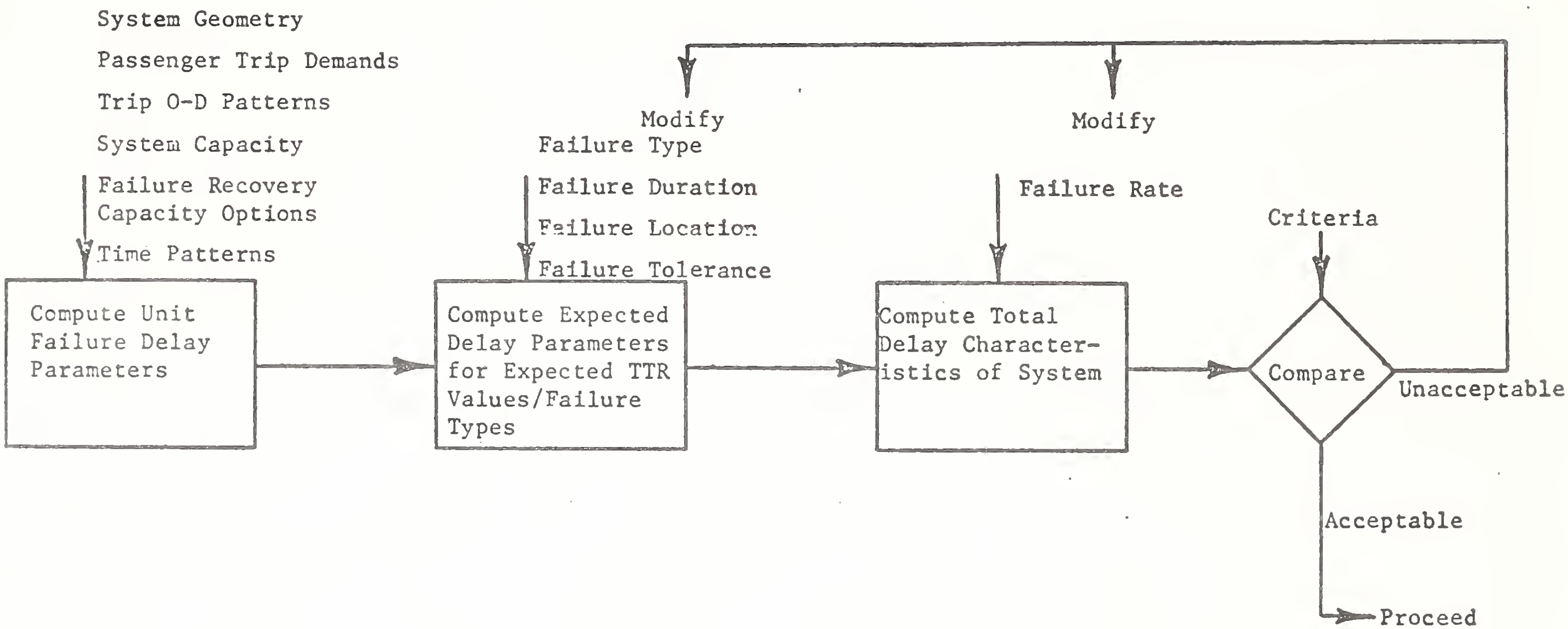


FIGURE B-2. PROCEDURE FOR DERIVING STATION DELAY PARAMETERS

Numerous system situations were examined to assess not only validity of this approach but also to define the specific procedures necessary to utilize it. The following subsection discusses the results of these analyses in the "building block" fashion alluded to earlier. The first examines the nature of the "unit failure" and characteristic response with fixed failure type and passenger demand variation. The subsequent subsection examines the effect of variations in these parameters. Subsequent to these subsections, techniques required to approximate expected system delay performance when numerous failure types, restore time, and failure rates are examined. The final subsection presents the resulting methodology.

B.2.1 Transit System Characteristic Response Function

As indicated earlier, the delay performance at a specific station can be approximated graphically, the geometry of which defines the delay parameters resulting from a failure of TTR duration. It was also pointed out that the shape and duration of the queue curves were the unknown variables. As illustrated in Appendix A, both of these variables are complicated functions not only of the delay dynamics of the station being analyzed but also those of other stations in the system. Therefore, the early examples analyzed in this program were directed toward defining relationships of these unknown variables with other known system parameters to enable graphics, such as Figure B-1, to be quickly generated.

B.2.1.1 Determination of SRT. With reference to Figure B-1, regardless of the shape of the queue dissipation curve, it follows that SRT is given by

$$SRT = \frac{DR \cdot TTR}{EC_a} \qquad \text{(B-1)}$$

where: SRT = Service restore time (time to dissipate queue to zero, restoring service demands at station to normal levels)

DR = Normal trip demand rate originating at station being analyzed

TTR = Time to restore equipment to operating condition

EC_a = Average excess capacity available for queue dissipation at station being analyzed.

As illustrated in Appendix A, the average excess capacity available to dissipate queues at any given station is less than the normal excess capacity at that station due to competition from upstream stations. When it is assumed that normal passenger ratios remain constant; that is, the normal excess capacity available at a station is shared by through passengers and originating passengers in proportion to their normal volumes, the following relationship results:

$$EC_a \approx \frac{DR}{LF} \cdot EC_n \tag{B-2}$$

where: DR = Normal trip demand rate originating at station being analyzed

LR = Normal link flow rate downstream from station being analyzed

EC_n = Normal excess capacity at station.

If the excess capacity for queue dissipation is derived solely from vehicle capacity considerations, Equation (B-2) becomes

$$EC_a \approx DR \left[\frac{1 - LF}{LF}\right] \tag{B-3}$$

where: LF = The load factor of vehicles leaving the station during normal conditions (passengers on board divided by vehicle capacity).

Similarly, Equation (B-1) becomes

$$SRT \approx \left[\frac{LF}{1 - LF}\right] \cdot TTR \tag{B-4}$$

Equation (B-4) is an approximation. The exact formulation for SRT would require the simultaneous solution of multiple, connected queuing problems (at each station). Such a process approaches mathematical intractability. The form of Equation (B-4) is simple and incorporates a variable (LF) which must be estimated to define required vehicle size during system design. Furthermore, in all examples examined in this program, which utilized actual simulation to define the queue dissipation curve, the resulting SRT values matched those predicted by Equation (B-4) very closely. It can be observed that Equation (B-4) exhibits behavior consistent with logic. For example, if the normal load factor is 1, there is no excess capacity available at that

station; hence, SRT is undefined and the queue developed during a failure cannot be dissipated. If the normal load factor is very small, the excess capacity is large. Hence, the queue developed during a failure can be dissipated quickly as Equation (B-4) would predict.

If the system has the capability to increase vehicle operating velocity as a method of recovering from a failure, Equation (B-2) becomes

$$EC_a \approx \frac{DR}{LF}\left[\frac{V_r}{V_n} - LF\right] \tag{B-5}$$

where: V_r = Vehicle velocity during the service restoration

V_n = Normal vehicle velocity

and Equation (1) becomes

$$SRT \approx \left[\frac{LF}{\frac{V_r}{V_n} - LF}\right] \cdot TTR\ . \tag{B-6}$$

The net effect of increasing vehicle velocity is a decrease in vehicle headways--in terms of time. The same effect can be achieved by decreasing vehicle spacing by inserting more vehicles into the system. With the scheme, Equation (B-2) becomes

$$EC_a \approx \frac{DR}{LF}\left[\frac{N_r}{N_n} - LF\right] \tag{B-7}$$

where: N_r = Number of vehicles during recovery time

N_v = Number of vehicles during normal operations

and Equation (1) becomes

$$SRT \approx \left[\frac{LF}{\frac{N_r}{N_n} - LF}\right] \cdot TTR\ . \tag{B-8}$$

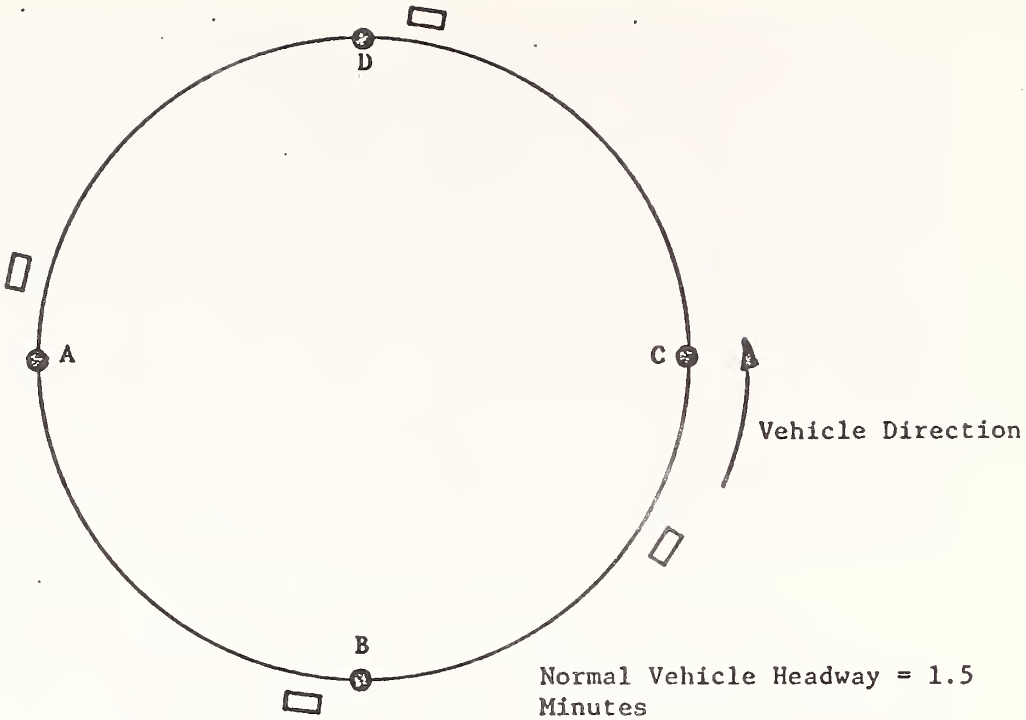
In many systems, these latter two options may not exist, either by design (velocity limits) or practical constraints (unavailability of extra vehicles or the inability to quickly dispatch and insert extra vehicles as required to recover from a failure). Hence, from a delay propensity standpoint in most systems, vehicle size will be the most influential parameter.

The above relationships serve to define SRT which, in turn, define a point in time following a failure when the delayed passenger queue is reduced to zero and the effect of the failures are no longer apparent to a new passenger entering the station being analyzed. Knowing this allows the number of passengers involved in the delayed queue to be estimated as will be discussed in a subsequent section. To determine the delay imposed on these passengers, however, the shape of the dissipation curve must be known.

B.2.1.2 Estimation of Queue Dissipation Curve Shape. The estimation of SRT utilizes an average value for excess capacity over the duration of the queue dissipation process. At any point in time during this process, however, the actual excess capacity at the station being analyzed can vary considerably from this average value. Like the average excess capacity, these intermediate excess capacity values are complex functions of queue dissipation dynamics interacting among stations. However, knowing SRT, at all stations from some value of TTR, reasonable approximations for the delay curves can be generated.

This process can best be illustrated through the use of an example. This example will also serve to illustrate the concepts discussed in the previous section. Figure B-3 illustrates the system used in this example. This system incorporates features such that any failure effects the entire system. As illustrated, four vehicles are used and indexed between stations.

The link loading data in Figure B-3 shows a maximum vehicle loading under normal conditions to be 25 passengers. This occurs on Links C-D and D-A. To provide excess capacity, a vehicle maximum loading value of 30 passengers is used.



Origin-Destination Matrix - Passengers Per Normal Vehicle Headway Interval

From	To	
A	B	4
A	C	4
A	D	2
B	C	2
B	D	2
B	A	8
C	D	1
C	A	10
C	B	2
D	A	2
D	B	2
D	C	1

Link Loads

A	B	13
B	C	19
C	D	25
D	A	25

FIGURE B-3. EXAMPLE TRANSIT SYSTEM

Figure B-4 illustrates the delay envelopes for each station in this hypothetical system. Station "C" is highlighted for this discussion. As noted, the delay envelope is generated in three steps.

(1) The initial rise of the delay envelope has a slope equal to the demand rate at that station and exists for a duration of TTR--in this case, two normal vehicle headway intervals.

(2) SRT is computed using appropriate equations from Section 5.2.1.1. In this case, excess capacity is derived solely from vehicle size. Therefore, Equation (4) is used, which, with the data provided in Figure 5-3 yields the following value for SRT for Station C of 5 TTR. This point is located on the delay envelope axis.

(3) The points defined by (1) and (2) are connected to complete the diagram. This connection process itself involves two steps. As noted previously, demands at each station compete for the normal excess capacity of a system. This competition is most fierce between adjacent stations. For example, with the four stations of the example, the queue dissipation capability at Station C is most influenced by the queue at Station B. It is influenced to a lesser extent by the queue at Station A because Station B, an intermediate station, acts as a buffer--some passengers from A exit at Station B. Therefore, as a first approximation, it is assumed that once the queue at Station B is dissipated, it is no longer competing for the system excess capacity. Therefore, after the queue at B is dissipated, the average excess capacity at Station C is the normal excess capacity (EC_n) which would be available under normal situations. (From Figure B-3, EC_n at Station C equals 5 passengers per normal vehicle headway interval.) Hence, at Station C, the last part of the delay curve can be approximated. By connecting these segments, a reasonable approximation to the excess queue curve is developed. Similarly, Station B performance can be developed. Stations A and D exhibit a characteristic triangular shape because the SRT values at these stations are less than those at the immediate upstream stations.

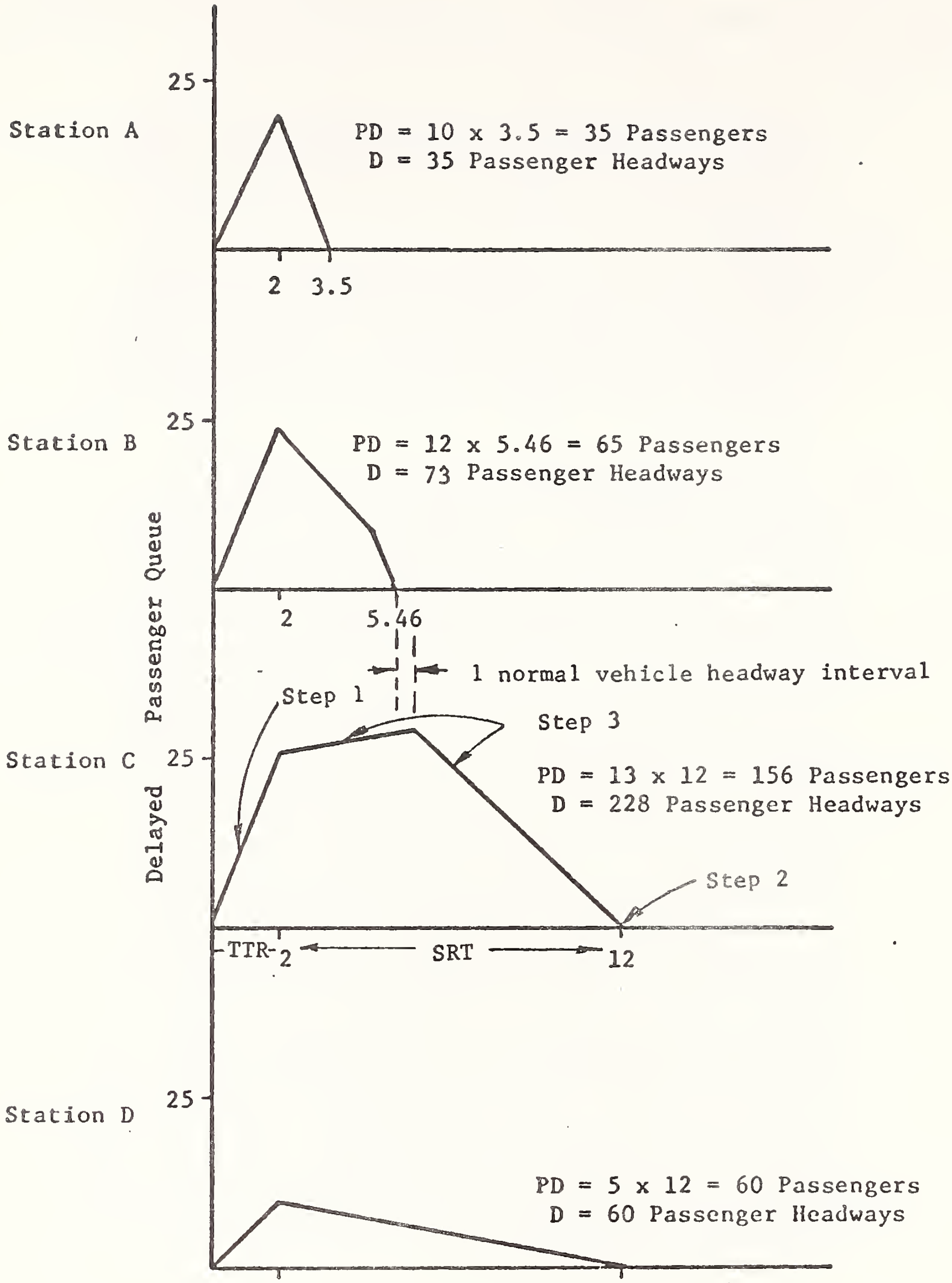


FIGURE B-4. DELAY ENVELOPES FOR FULL-STOP FAILURE

Having these shapes, one can compute the total passenger-minutes delay resulting from the postulated failure, as well as the number of passengers delayed.

Exercising this process for some arbitrarily selected calue of TTR, denoted by TTR*, the number of passengers delayed at station, and the cumulative duration of the delay can be estimated. The delay parameters are denoted PD* and D*, respectively. The unit failure response, in the form of these values, can be used to generate delay parameters for many expected failure situations by applying scaling procedures.

B.2.2 Response to Full System Stoppage Failures

Full system stoppage failures have the same character of the unit failure with the exception that the TTR value is taken to represent the actual values expected rather than an arbitrary TTR*. As discussed previously, the number of passengers delayed at stations by a full stop failure proportional to TTR while the cumulative delay is proportional to TTR^2. Therefore, to estimate the delay parameters for a number of full stop failures, the following relationships apply:

$$PD = PD^* \frac{\overline{TTR}}{TTR^*} \tag{B-9}$$

where: PD = Total number of passengers delayed per average failure

PD* = Total number of passengers delayed per unit failure

$\overline{TTR}$ = Mean time to restore for full-stop failures

TTR* = Unit failure downtime.

The relationship for estimating the cumulative delay (D) experienced by delayed passengers during an average failure is

$$D = D^* \frac{\widetilde{TTR}^2}{TTR^{*2}} \qquad \text{(B-10)}^{(a)}$$

where: D = Cumulative delay experienced by passengers per average failure

D^* = Cumulative delay per unit failure

$\widetilde{TTR} = \sqrt{\frac{\sum_i \lambda_i \overline{TTR}_i^2}{\sum_i \lambda_i}}$, RMS value of $\overline{TTR}_i$ for failure mix

TTR^* = Unit failure downtime.

The probability of delay can be estimated by the following relationship:

$$Pr_d = \frac{PD^*}{DR} \frac{\overline{TTR}\,\lambda}{TTR^*} \qquad \text{(B-11)}^{(b)}$$

where Pr_d = Probability of being delayed at station on an average trip

DR = Total system trip demand rate

λ = Failure rate.

(Other variables defined as before.)

Expected delay can be estimated by the following relationship:

$$ED = \frac{D^*}{DR} \frac{\widetilde{TTR}^2 \lambda}{TTR^{*2}} \qquad \text{(B-12)}$$

where ED = Average station delay expected on average trip.

(Other variables defined as before.)

Care should be exercised in using these relationships to ensure consistency of units. A new term has been introduced, $\widetilde{TTR}$, in the delay scaling procedures. Because delay (the <u>area</u> under the delayed queue curve

---

(a) The eauations in this section are all relevant only to full system stoppage failures (the subsection title). Hence, all parameters (excluding unit failure parameters) carry an implicit subscript denoting this relevance. In this document, these subscripts are not included to enhance clarity. This philosophy is continued in all subsequent subsections.

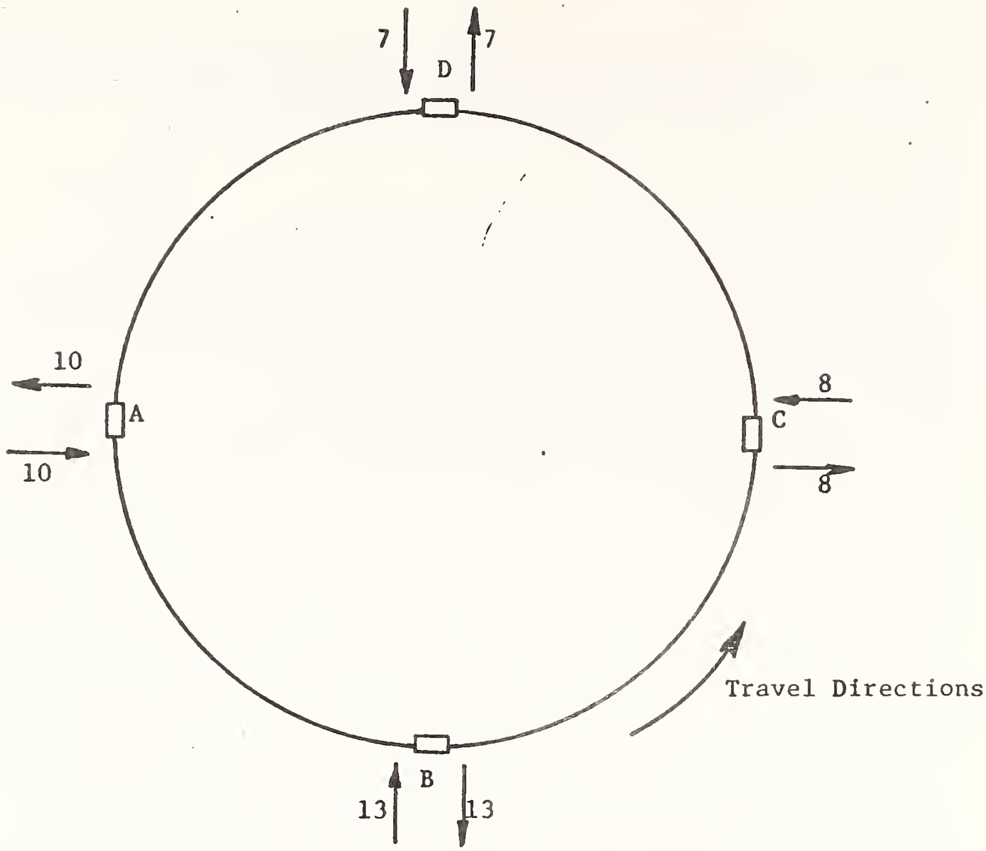
(b) In this formulation, the assumption is made that $\lambda \overline{TTR} << 1$. This is reasonable based on sparse data from existing systems which place values for both of these parameters within an 0.1 magnitude. If this assumption is thought to be too liberal, λ may be replaced with $\lambda/(1 + \lambda \overline{TTR})$. This assumption is used generally throughout this document.

of Figure 1) is proportional to TTR^2, a quadratic mean (or RMS value) is required to scale unit failure response (D*) properly.

B.2.3 The Effects of Headway Closing Options

The first point examined was the failure mode in which one vehicle stops and others are permitted to close in behind (utilizing the "cushion" referred to in many analyses). This operational mode is assumed in most studies which focus on individual passenger expectation. In these studies, a downstream vehicle failure may not affect the delay experienced in boarding a system or en route if the failed vehicle becomes operational before safe headway constraints are met. Interpreting such probabilities in terms of expected system delay parameters implies an assumption of an unlimited source of vehicles. In a real system, there is a very limited source of vehicles available. Furthermore, if the vehicles "bunch", as they would if the cushion is utilized, they must be spread out quickly to eliminate delays encountered as the result of the gap existing elsewhere in the system. A station would see several vehicles at closer-than-normal headways, then a gap would appear--inducing delays at that station.

Essentially, the appearance of this gap represents another "failure" insofar as service to a station is concerned. It is important to get the system back on normal speed/headway relationships to minimize this recurring failure. If the failure management strategy is to create normal headways immediately following the failure by holding trailing vehicles as necessary, the delay curves at stations are nearly identical to those resulting from full-stop failure except that they do not occur at all stations at the same time. Rather, they are displaced, in time, from one another. If the failure management strategy incorporates some programmed antibunching sequence to restore normal headways following a failure, the delays incurred may vary considerably from the full-stop failure effects. As an example to illustrate this, the system utilized in Appendix A was extended as illustrated in Figure B-5. Figure B-6 shows a train graph of normal operation, a failure of 4 normal headways duration occurring between Stations A and B, and a failure management strategy which requires establishing a temporary headway of 1/2 the normal headway immediately following the failure correction. This headway is maintained for one full cycle of this system, at which time headways are readjusted to achieve normal values.



Trip Matrix (Trips Per Vehicle)

From	To	No.
A	B	5
A	C	3
A	D	2
B	C	4
B	D	4
B	A	5
C	D	1
C	A	3
C	B	4
D	A	2
D	B	4
D	C	1

Other Assumptions

1. System consists of 4 vehicles.
2. Vehicles are temporally spaced so that each enters a station on a fixed schedule.
3. There are no station bypass options.
4. A failure anywhere in the system causes a complete system shutdown.

Station Flow Parameters

Station	On	Off	Thru
A	10	10	9
B	13	13	6
C	8	8	11
D	7	7	12

Link Loading Parameters

Link	Passengers
A-B	19
B-C	19
C-D	19
D-A	19

FIGURE B-5. EXAMPLE SYSTEM FOR ESTIMATING EFFECTS OF "CUSHION" UTILIZATION

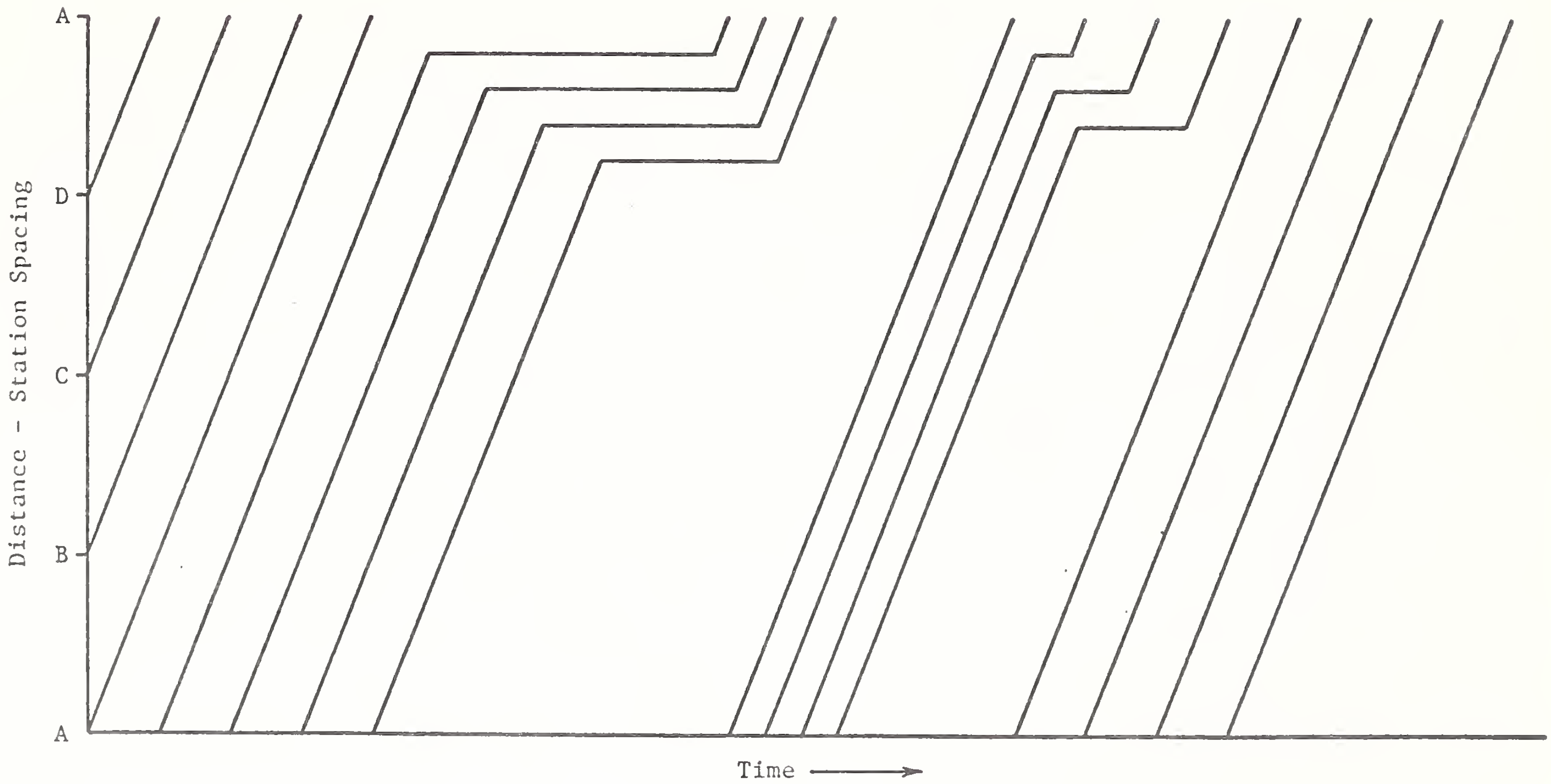


FIGURE B-6. TRAIN GRAPH OF EXAMPLE FAILURE SHOWING CUSHION UTILIZATION AND HEADWAY RECOVERY

Figure B-7 shows the resulting station delay curves for this example. The solid curves were determined from simulation results while the dotted lines represent the curves which would have been predicted using the full-stop relationships discussed in Section B.2.1. As can be seen, the actual delays encountered are greater than those which would have been encountered if the system had been totally stopped. If the failure management strategy had permitted the vehicles to travel bunched for a longer period of time, the differences would have been greater.

Who, then, benefits from cushion utilization? Three groups of passengers.

(1) Passengers who are on board vehicles which trail a failed vehicle and who are destined for a station on the upstream side of the failure. If the headway compression is sufficient, they may reach their destination with no delay. If the failure is cleared quick enough, trailing vehicles may not need to stop. Hence, no passengers on board would see a delay.

(2) Passengers at stations upstream from the failure who are destined to stations between them and the failure and who are fortunate enough to catch a trailing vehicle prior to the appearance of the gap. They may complete their trip without a dealy.

(3) Passengers at stations immediately upstream from the failure. After equipment operation is restored, a rapid succession of vehicles will be available. This could reduce the average delay experienced by those passengers which are fortunate enough to catch the first wave of vehicles following the failure.

Therefore, while certain passengers may receive some benefit from headway compression, such benefits accrue to the detriment of other passengers.

Headway compression can become a beneficial mode of failure recovery if the lead vehicles in the bunch have an increased speed capability, permitting normal headways to be established without slowing down trailing vehicles. Also, if the option exists for inserting spare vehicles into the system to fill the gap would be useful. However, in any practical system, neither of these approaches is likely. Therefore, the conclusion to be drawn is that, in terms of delay effects, headway compression capability is not an asset.

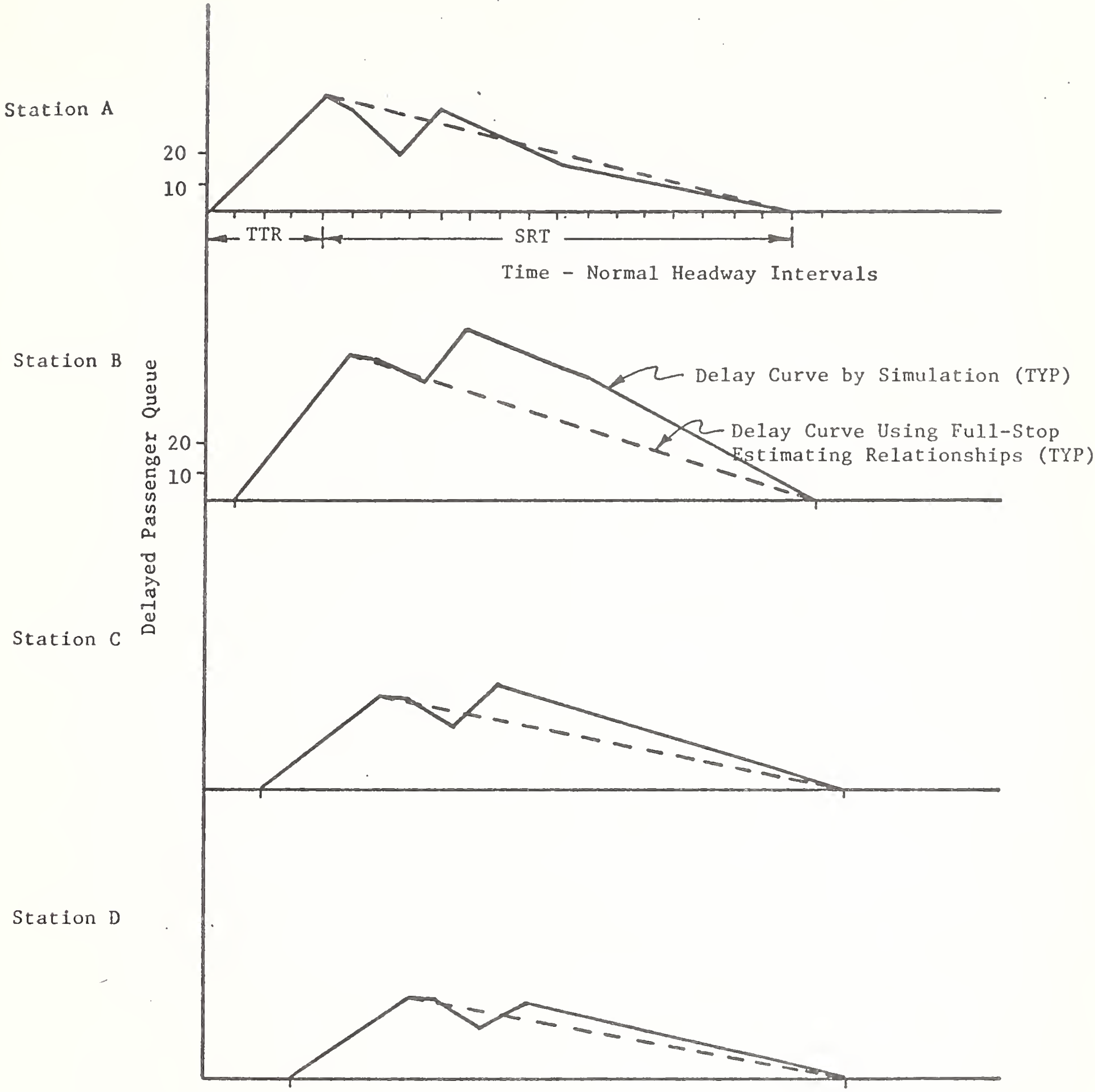


FIGURE B-7. DELAY ENVELOPES OF EXAMPLE FAILURE WITH CUSHION UTILIZATION AND HEADWAY RECOVERY .

Rather, it may be a liability. While we have not examined all situations, we have examined enough examples to support this thesis for the type of system depicted in Figure B-5, a totally connected system with reasonably balanced station demands and short headways such that failure duration is long with respect to the normal vehicle headway. Irrespective of this, however, permitting motion of vehicles not directly affected by the failure would have positive psycological benefits on the passengers. These benefits accrue to passengers en route. From the standpoint of station delays, the delay parameters derived for a full-stop failure are appropriate.

B.2.4 Effects of Increased Headway Failures

Failures which result in reduced operating speed or a reduction in the number of operating vehicles induce the same effects in terms of delays incurred at stations. In both cases, the apparent effect is a decrease in service frequency or an increase in operational headway.

In the case of reduced velocity failures, numerous examples were investigated which led to an observation relating the type of failure to the characteristic response discussed in Section B.2.1. Figure B-8 summarizes the results. In this figure, the solid curve represents the delay envelope for a failure which results in a 50 percent speed reduction. The dotted curve represents the delay envelope for a full-stop failure--a hypothetical failure which has a queue dissipation curve coincident with that of the slow-speed failure, i.e., TTR + SRT for this hypothetical full-stop failure equals (TTR + SRT) for the slow-speed failure. As can be seen, the slow-speed failure can be considered to result in some portion of a full-stop failure with a TTR of 1/2 the slow-speed TTR. If the sawtooth effect is neglected, the slow-speed delay envelope is exactly 1/2 of the depicted full-stop delay envelope. These relationships can be generalized as follows:

$$TTR_{fs} = TTR_{ss} \left[1 - \frac{V_f}{V_n} \right], \qquad \text{(B-13)}$$

where TTR_{fs} = TTR for full-stop failure which shares the delay envelope of the slow-speed failure,

TTR_{ss} = Duration of the slow-speed failure,

V_f = Vehicle velocity during the failure,

V_n = Normal vehicle velocity.

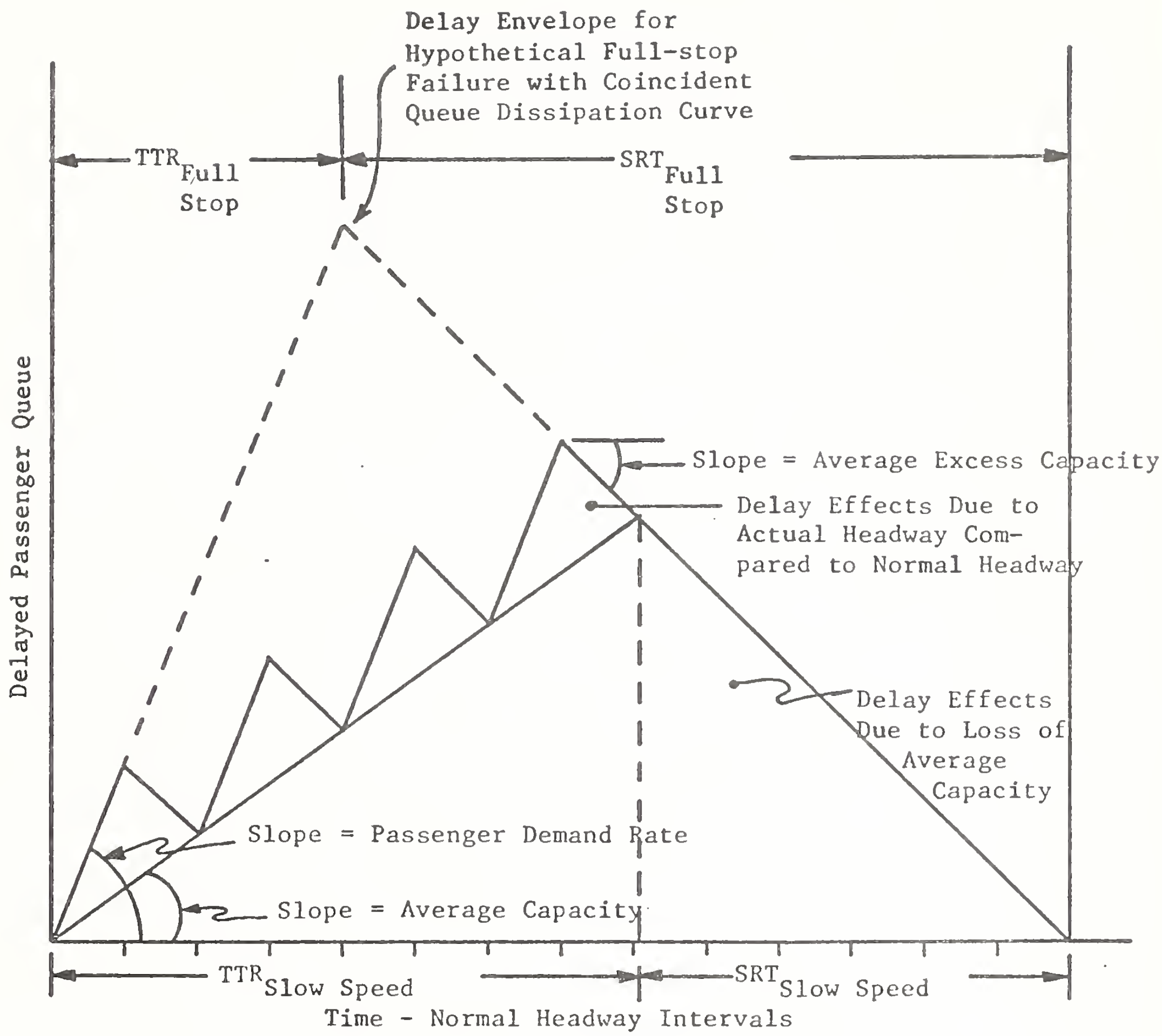


FIGURE B-8. SLOW-SPEED FAILURE DELAY ENVELOPE

The cummulative delay relationship is

$$D_{ss} = D_{fs} \left[1 - \frac{V_f/V_n}{\left(\frac{LF}{1 - LF}\right)\left(1 - \frac{V_f}{V_n}\right)} \right] \tag{B-14}$$

where D_{ss} = The cumulative passenger delay resulting from the slow-speed failure (If the values in brackets are negative, D_{ss} = 0.).

D_{fs} = Cumulative passenger delay incurred in the overlapping full-stop failure.

These relationships can be used to translate the characteristic response of the system into cumulative delay expectations for any slow-speed failure of any duration. By similarity, Equations (B-13) and (B-14) can be used for failures which require a temporary loss of one or more vehicles from service by substituting N_f/N_n for V_f/V_n,

where N_f = Number of vehicles operating during failure,

N_n = Number of vehicles required for normal service frequencies.

In both cases, the "sawtooth" portion of the delay envelope has been neglected. This periodic effect is due to the erratic schedule during the failure. As indicated in Appendix A, while this recurring, short-duration delay effect is present, the real problem associated with these types of failures is the loss of capacity which accompanies them. Hence, neglecting the sawtooth effect may be justified. If not, an additional term must be added to Equation (B-14).

To determine the number of passengers delayed during failures of this type, simple methods are not available. Neglecting again the sawtooth effect, the ratio of passengers delayed due to the slow-speed failure and the passengers which would be delayed due to the hypothetical full-stop failure is a function of the duration of the slow-speed failure. Unfortunately, this function is difficult to define mathematically. As TTR becomes large, the ratio of passengers delayed due to the slow-speed failure to those passengers which would have been delayed due to the hypothetical full stop failure approaches unity. For estimating purposes, therefore, one can assume that, for those stations where station delays accrue, all passengers entering that

station during the failure and its associated queue dissipation time will be delayed. If more precision is required, graphic techniques would be required.

Using Equations (B-13) and (B-14) and the above assumption, the following relationships apply for estimating delay parameters for increased headway failures:

$$D_j = \frac{\overline{TTR}_{ss}}{TTR^*} D_j^* \left(\frac{LF_j - V_f/V_n}{LF_j} \right) \qquad \text{(B-15)}$$

where: D_j = Cumulative delay incurred at station j due to slow-speed failure

$\overline{TTR}_{ss}$ = Duration of slow-speed failure

TTR^* = Duration of unit failure

D_j^* = Cumulative delay incurred by passenger at station j due to unit failure

LF_j = Normal load factor of link leaving station j

V_f = Vehicle velocity during failure

V_n = Normal vehicle velocity.

Equation (13) is constrained as follows:

If $V_f/V_n \geqq LF_j$, $D_j = 0$.

It follows, therefore, that the number of passengers delayed is accumulated only at those stations where delay is accumulated. the relationship governing the number of passengers delayed is:

$$PD_j = \frac{\overline{TTR}_{ss}}{TTR^*} \left[1 - \frac{V_f}{V_n} \right] PD_j^* \qquad \text{(B-16)}$$

where: PD_j = Number of passengers delayed at station j due to queue buildup at stations under a slow-speed failure

PD_j^* = Number of passengers delayed at station j due to unit failure.

(Other variables defined as before.)

In this equation, j takes only values permitted by the constraint imposed on Equation (B-15).

The relationships governing probability of delay and expected delay are as follows:

$$Pr_d = \frac{\sum PD_j}{DR} \lambda_{ss} \tag{B-17}$$

where: Pr_d = Probability of incurring a delay on a random trip due to a slow-speed failure

λ_{ss} = Failure rate for slow-speed failure.

(Other variables defined as before.)

The relationship for estimating expected delay is

$$ED = \frac{\sum D_j}{DR} \lambda_{ss} \tag{B-18}$$

where: ED = Average station delay per average trip due to slow-speed failure.

(Other variables defined as before.)

Relationships governing the vehicle out-of-service failure are as follows:

$$D_j = \frac{\overline{TTR}_{vos}}{TTR^*} D_j^* \left(\frac{LF_j - N_f/N_n}{LF_j} \right) \tag{B-19}$$

where: D_j = Cumulative delay at station j due to vehicle-out-of-service failure

$\overline{TTR}_{vos}$ = Duration of VOS failure

N_f = Number of vehicles operating during failure

N_n = Number of vehicles normally operating.

(Other variables defined as before.)

The number of passengers delayed may be estimated by the following relationship:

$$PD_j = \frac{\overline{TTR}_{vos}}{TTR^*} \left[1 - \frac{N_f}{N_n} \right] PD_j^* \tag{B-20}$$

where: PD_j = Number of passengers incurring delays at station j due to VOS failure.

(Other variables defined as before.)

Equations (B-19) and (B-20) are constrained as follows:

If $N_f/N_n \geqq LF_j$, then $D_j = 0$ and $PD_j = 0$.

The probability of delay can be estimated by the following relationship:

$$Pr_d = \frac{\sum PD_j}{DR} \lambda_{vos} \qquad \text{(B-21)}$$

where: Pr_d = Probability of incurring a delay on an average trip due to VOS failure

λ_{vos} = Failure rate for VOS failure.

(Other variables defined as before.)

The following relationship can be used to estimate expected delay:

$$ED = \frac{\sum D_j}{DR} \lambda_{vos} \qquad \text{(B-22)}$$

where: ED = Average station delay per average trip due to VOS failures.

(Other variables defined as before.)

B.2.5 Effects of Partial System Failures

The previous types of failures all apply uniformly to complete systems, that is, the failure affects the entire system. Such would be the case with simple circulation systems with on-line stations, no passing capability, no turn-around capability, and no reverse-running capability. For these types of systems, a failure cannot be isolated. For systems where a failure can be "disconnected", allowing the remainder of the system to operate, different failure effects accrue. Only passengers which require use of the failed portion will be delayed. Station queues will develop involving all trips which are affected by the partial system failure. After the failure is removed, these queues will dissipate according to the average excess capacity available at those stations. If the excess capacity is very large relative to the demands, such that these delay queues will dissipate immediately, a relationship exists to the delay parameters computed for a full-stop failure of equal duration. In this situation, the delays incurred at any station are related to those which would result from a full-stop failure by the following relationship:

$$PD_{pf} = PD_{fs} \cdot \frac{T_a}{T_t} \tag{B-23}$$

where PD_{pf} = Number of passengers delayed at a station due to a partial system failure of TTR duration

PD_{fs} = Number of passengers delayed at that station due to a full-stop failure of equal TTR

T_a = Passenger trips originating at station during the failure which require the use of the failed portion of the system

T_t = Passenger total trips generated at that station during the failure.

If the excess capacity is not large, interaction among station queues again complicates the service restoration process, resulting in deviations from the above relationship. Equation (B-23) can estimate above or below simulation results. Fortunately, when all possible partial failures are analyzed for any system, the "over estimates" tend to balance the "under estimates". Therefore, Equation (B-23) can be considered to represent a reasonable approximating relationship.

Using this relationship, the estimation of delay parameters based on unit failure response is accomplished as follows:

By generalizing this relationship to the full system and using the unit failure as the full-stop failure, estimates of delay parameters based on unit failure response is accomplished. The expression for number of passengers delayed is

$$PD_k = \frac{PD^*}{TTR^*} \overline{TTR}_k \frac{T_k}{T_t} \tag{B-24}$$

where: PD_k = Number of passengers delayed due to a failure which closes the kth portion of the system

PD^* = Number of passengers delayed per unit failure

TTR^* = Duration of unit failure

$\overline{TTR}_k$ = Mean-time-to-restore failure which denies use of the k portion of system

T_k = The number of trips (originating anywhere in the system) which require the use of the kth portion of the system during some time interval

T_t = The total number of trips generated during this time interval.

The cumulative delay experienced by the passenger can be estimated by the following relationship:

$$D_k = \frac{D^*}{TTR^{*2}} \widetilde{TTR}_k^2 \frac{T_k}{T_t} \tag{B-25}$$

where: D_k = Cumulative delay of passengers experiencing station delays due to failures of kth portion of the system

D^* = Cumulative delay experienced by passengers due to unit failure

$\widetilde{TTR}_k$ = RMS value of TTRs for failures affecting kth part of the system

(Other variables defined as before.)

By summing these relationships over k, general delay statistics can be derived. The expression for estimating probability of delay is

$$Pr_d = \frac{\sum(PD_k \lambda_k)}{DR} \tag{B-26}$$

where: Pr_d = Probability of delay on average trip due to partial system failure

λ_k = Failure rate for failures affecting kth part of the system

(Other variables defined as before.)

Expected delay can be estimated by the following relationship:

$$ED = \frac{\sum D_k \lambda_k}{DR} \tag{B-27}$$

where: ED = Average delay expected for an average trip due to partial system failures

λ_k = Failure rate for failures affecting kth part of the system.

(Other variables defined as before.)

B.2.6 Effects of Varying Passenger Demands

In the previous section, relationships have been developed whereby the delay effects of many different failure situations can be scaled from estimates developed on the basis of some full-stop failure of arbitrary duration. These latter delay parameters are relatively simple to compute and can be interpreted as the propensity of a unit of system downtime to induce passenger delays. However, these values are dependent on the trip demands and the O-D patterns existing at the time of the failure.

Therefore, to be useful in a real sense, the implications of varying demands such as would occur in a real situation must be comprehended. As demand patterns at stations vary throughout a typical day, values for both demand rate and average excess capacity vary. Unit failure response can vary considerably. Furthermore, the influence of demand levels is not linear. If the excess capacity in the system is small, the delay parameters are more nearly proportional to the square of the demand level. Hence, any specific unit failure response under some demand pattern can be considered to apply only to that time period over which that demand pattern exists. Comprehending all variations in demand patterns which may exist during the operating life of a transit system is an impossible feat. Therefore, some selected set of patterns must be selected which represent typical patterns. In selecting these "typical" patterns, emphasis should be placed on matching peak demands.

Off-peak demands can be averaged with much less of an impact on overall system delay performance. Therefore, as an approximating technique, for systems which exhibit peaks in demand (as, for example, would be the case in any system serving commuter traffic), the daily demand patterns can be approximated by three average values.

(1) The morning peak average over some period, depending on the duration of the peak

(2) The afternoon peak, averaged over the duration of the peak

(3) The off-peak average demands.

B.2.7 Effects of Multiple Failures

The previous sections have not dealt with the concept of multiple failures. Each failure is assumed to be cleared well in advance of another. The principle enhances the potential for estimating delay effects for actual expected failures by scaling the results of unit failure delay parameters because a "normal" condition exists prior to the failure in both cases. Certain failures cannot be treated this way, however, because the initial system condition is in an abnormal state due to some preceding failure. Take, for example, the case of a vehicle-out-of-service failure. It is likely that such a failure will involve

(1) A full-stop failure of sufficient duration to remove the vehicle (e.g., pushing to a siding or station for removal)

(2) A period following removal of sufficient duration to prepare, check out, and insert a substitute vehicle into service.

Each of these can be considered a separate failure with the initial state of the second failure determined by the system response to the first failure. For this situation, scaling laws cannot be derived. Superposition is not appropriate either. Graphical techniques can still be utilized to handle these special cases, however. Such a process is illustrated in Figure B-9. This depicts station delays accruing from a vehicle failure which require removal from service. It takes five normal vehicle headway intervals to clear the system for operation, followed by a period of 20 normal headway intervals to insert a substitute vehicle into the system.

The initial part of the delay envelope is identical with that which would result had the full-stop failure occurred by itself. With normal equipment operation following the full-stop failure, an average excess capacity would exist, shown on Slope A in Figure B-9.

However, because of the ensuing vehicle reduction, this excess capacity is not available; hence, recovery is hampered. The resulting queue dissipation curve has a slope which can be computed as follows:

Modified excess capacity = Slope B

$$= (1 + EC_n)\frac{N_f}{N_n} - 1 , \qquad \text{(B-28)}$$

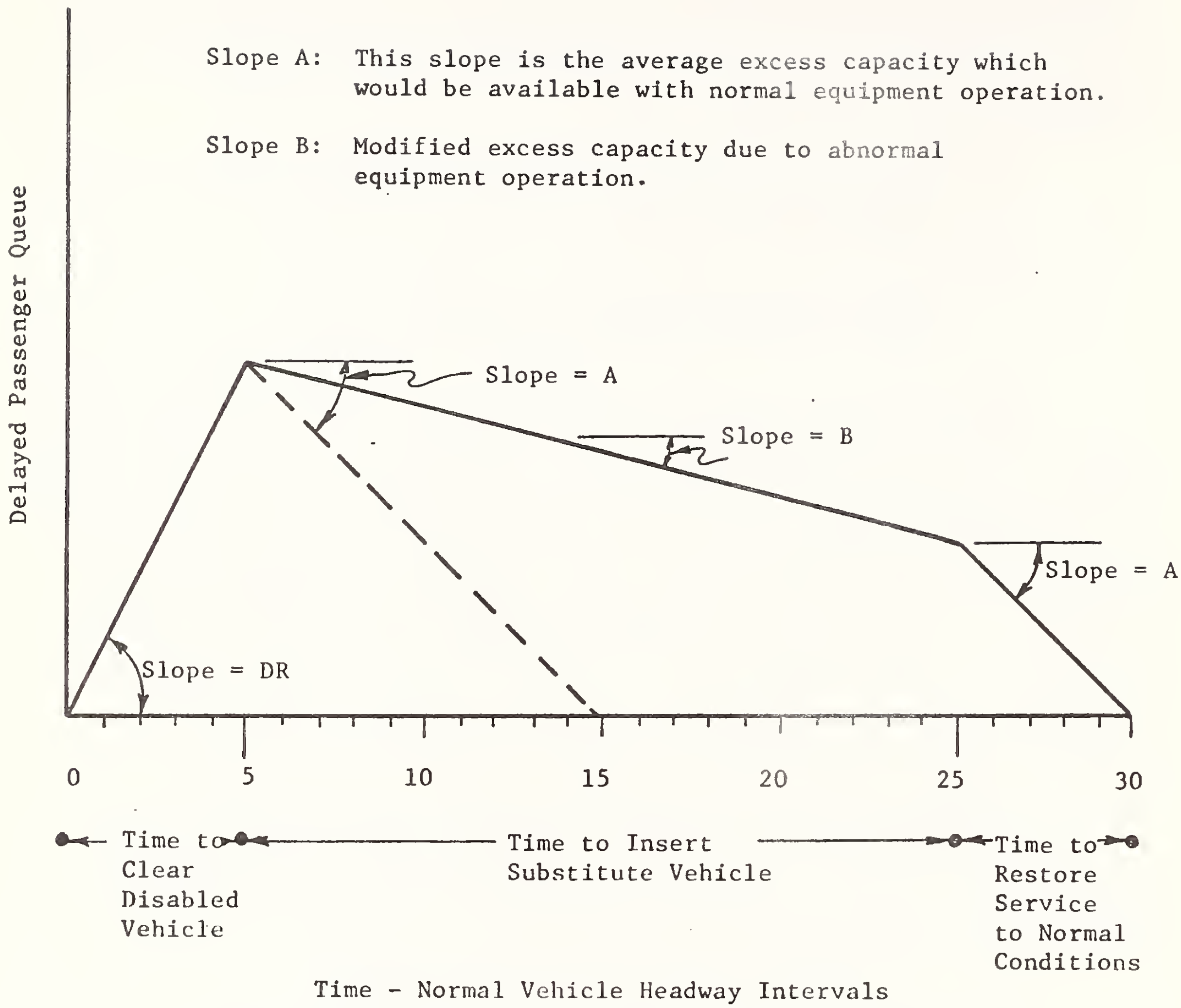


FIGURE B-9. EXAMPLE OF MULTIPLE FAILURE DELAY PROFILE

where EC_n = Average excess capacity available with normal system operation,

N_f = Number of vehicles in service during off-normal period,

N_n = Number of vehicles normally available.

This slope is maintained until the substitute vehicle is in service, at which point the EC_n is available to rapidly dissipate the remaining queue.

When system alternatives have such a compound failure mode, estimated of delay parameters must be done using such techniques.

B.3 EN ROUTE DELAYS

The entire discussion thus far has dealt with techniques for estimating delays which accrue at stations due to a failure in the system. The other component of delay is that which accrues to passengers en route at the time of the failure. Estimating this effect is considerably less difficult to handle because queue dynamics are not influencing the results.

Classical probability mathematics as applied in the literature are applicable. However, these techniques can become extremely complex in estimating en-route delay potential for the variety of failure types and system configurations possible even for simple systems. As an alternative, en-route delay potential can be estimated using techniques which parallel those above for estimating station delay potential. In most cases, the data upon which such estimates are based will have been generated in the station delay calculations. In general, the procedure is to simply "observe" the effect of various types of failures as might be expected in any specific system situation and deduce expected system performance from these observations.

B.3.1 Response to Full System Stoppage Failures

The full-stop failure type is the simplest to deal with. Each time a failure occurs, all passengers on board vehicles at that time are delayed. Using average values, this amounts to the average instantaneous link loading of the system. The duration of the delay is equal to the time to restore the equipment to operating conditions.

Hence, for each failure, the number of passengers delayed en route is

$$PD = N_d C_v \overline{LF} \tag{B-29}$$

where: PD = Average number of passengers delayed en route for a full stoppage failure

N_d = Average number of vehicles deployed between stations

C_v = Vehicle capacity

$\overline{LF}$ = Average load factor for all vehicles operating.

Cumulative delay can be estimated by the following relationship:

$$\overline{D} = \overline{TTR} \tag{B-30}$$

where: $\overline{D}$ = Average delay per delayed passenger due to full stoppage failure

$\overline{TTR}$ = Mean time to restore equipment to normal operating levels.

The probability of delay can be estimated by the following relationship:

$$Pr_d = \frac{N_d C_v \overline{LF} \lambda}{DR} \tag{B-31}$$

where: Pr_d = Probability of being delayed en route average trip

λ = Failure rate for full system stoppage failure

(Other variables defined as before.)

Expected delay can be estimated by the following relationship:

$$ED = \frac{N_d C_v \overline{LF} \lambda \overline{TTR}}{DR} \tag{B-32}$$

where: ED = Average delay expected on average trip due to full system stoppage failures

(Other variables defined as before.)

Because of the nature of this failure type, the en route delays will involve vehicle stoppages. Hence, the previous general delay parameters are also stoppage parameters.

B.3.2 Response to Stoppage Failure With Headway Closing Option

If the system allows vehicle headway closure, the estimating procedure can be quite difficult. The number of passengers delayed en route and their average delay can be manipulated by the particular antibunching scheme employed. Some passengers on board at the time of the failure may be delayed, while others who board after the failure has been corrected may be delayed. There is an infinity of trade-off options as to the number of passengers delayed and the duration of the delays. Additionally, station delays may substitute for en-route delays. In general, the benefits which accrue to any specific passenger are gained at the expense of other passengers. As discussed in Section B.2.3, the total amount of passenger-hours of delay is approximately constant. Therefore, to simplify the estimation procedure, treating the headway closing situation as a full-stop failure is a reasonable first-order approximation for general delay estimation.

For en-route stops, however, such liberties are not appropriate. Normally, vehicles operate at some operating headway (OH). They may close to some lower minimum headway (MH). MH may be defined by safety considerations or operating rules. The difference between these two headway values is termed cushion headway (CH) and represents the amount (in time units) by which one vehicle may close upon another. If a vehicle is stalled on the line, the immediate trailing vehicles will close in as permitted by CH. If the failed vehicle is down for a less than one cushion headway, the trailing vehicle will not be required to stop (although it may be slowed or delayed at a station to restore normal vehicle spacings). If the failed vehicle is stalled for a period of time greater than one cushion headway, trailing vehicles will be required to stop. The number of vehicles required to stop, including the failed vehicle, can be derived from the following expression:

$$N = \frac{\overline{TTR}}{CH}, \tag{B-33}$$

where N = Number of vehicles stopped as a result of the stoppage failure,

$\overline{TTR}$ = Average failure duration,

CH = Cushion headway.

In Equation (B-33), N should be taken as the next highest integer value of the fraction $\overline{TTR}/CH$. The maximum number of vehicles which can be stopped is limited by the number of vehicles in service. If this value is denoted N_n, Equation (B-33) is valid only when $\overline{TTR} \leqq N_n CH$. For $\overline{TTR} > N_n CH$, the number of vehicles delayed is equal to N_n.

The average delay per delayed vehicle is similarly a two-part formulation,

when $\overline{TTR} \leqq N\ CH$,

$$D_v = \frac{\overline{TTR}}{2} \tag{B-34}$$

where: D_v = Average delay per delayed vehicle.
(Other variables defined as before.)

When $\overline{TTR} > N\ CH$,

$$D_v = 3/2\ \overline{TTR} - N_n CH. \tag{B-35}$$

Passenger delay parameters can be derived from these relationships as follows:

When $TTR \leqq N_n CH$, the following relationships can be used to estimate the probability of delay and expected delay.

$$Pr_d = \frac{C_v \overline{LF}}{DR\ CH}\ \overline{TTR}\ \lambda \tag{B-36}$$

where: Pr_d = Probability of incurring a stoppage en route on an average trip

C_v = Vehicle capacity

$\overline{LF}$ = Average load factor of all vehicles

DR = Passenger demand rate

$\overline{TTR}$ = Average duration of failure

λ = Frequency of occurrence of failure.

$$ED = \frac{C_v \overline{LF}}{DR\ CH} \left(\frac{\overline{TTR}^2 \lambda}{2} \right) \tag{B-37}$$

where: ED = Average duration of en route stoppage expected on an average trip.

(Other variables defined as before.)

When TTR > N_nCH, the following relationships are applicable:

$$Pr_d = \frac{N_n C_v \overline{LF}}{DR} \lambda \qquad \text{(B-38)}$$

where: Pr_d = Probability of incurring a stoppage en route on an average trip,

and $$ED = \frac{N_n C_v \overline{LF} \lambda}{DR} (3/2\ TTR - N_n CH) \qquad \text{(B-39)}$$

where: ED = Average duration of en route stoppage expected on an average trip.

Many operating strategies may be used to limit the number of vehicles stopped en route (e.g., slowing down trailing vehicles or holding at upstream stations). In these situations, a link-by-link assessment is required to determine the number of vehicles actually involves in a stoppage incident. From this information, passenger delay parameters can be derived.

B.3.3 Response to Increased Headway Failures

During a failure which restricts vehicle velocity to some level below the normal velocity, all passengers who board vehicles during the failure duration ($\overline{TTR}$) will be delay ed due to the lower trip speed. (Some passengers who boarded just prior to the failure will experience delay; some passengers who board just prior to system recovery will not be delayed. These two effects are assum ed to cancel one another.) The rate of passengers boarding is not obvious. At stations where queues do not develop, the passenger boarding rate is identical to the normal demand rate (DR). At stations where queues do build up, the actual boarding rate is less than the normal demand rate. Discounting this difference, and assuming the normal demand rate, therefore, tends to overestimate passenger delay parameters. However, this is a conservative estimate and the relationships involved are simple. Therefore, the approximation is used. Hence,

$$PD = DR\ \overline{TTR}_{ss} \tag{B-40}$$

where: PD = Number of passengers delayed en route due to slow-speed failure

DR = Average passenger demand rate (This should actually be passenger boarding rate, which may be less than demand rate. However, using DR is reasonable approximation and errs on the conservative side.)

$\overline{TTR}_{ss}$ = Average duration of the slow-speed failure.

The cumulative delay may be estimated by the following relationship:

$$\overline{D} = TT\left(\frac{V_n}{V_f} - 1\right) \tag{B-41}$$

where: $\overline{D}$ = Average delay per delayed passenger due to slow-speed failure

TT = Trip time for average trip

V_n = Normal vehicle velocity

V_f = Vehicle velocity during failure.

The following relationships can be used to estimate the probability of delay and expected delay:

$$Pr_d = \overline{TTR}_{ss}\ \lambda_{ss} \tag{B-42}$$

where: Pr_d = Probability of being delayed on average trip due to slow-speed failure

λ_{ss} = Failure rate of slow-speed failures.

$$ED = \overline{TTR}\lambda_{ss}TT\left(\frac{V_n}{V_f} - 1\right) \tag{B-43}$$

where: ED = Average delay expected on average trip due to slow-speed failure.

(Other variables defined as before.)

Equation (41) estimates the probability of delay en route due to a slow-speed failure. As discussed in Section 5.2.4, some of the passengers

Because of the nature of the slow-speed failure, no stoppage incidents are encountered. Vehicle-out-of service failures do not impact en-route passengers except those aboard the failed vehicles. Hence, on a per failure basis, the number of passengers delayed is equal to the average number of passengers aboard an average vehicle. The duration of the delay incurred must be estimated from the operational procedures and time required to transfer these passengers to unfailed vehicles. The delay parameters must, therefore, be estimated with knowledge or assumptions regarding these procedures. No general formulation exists.

B.3.4 Response to Partial System Failures

When the system is configured such that a failure can be isolated to limit service on only a part of the system, en-route delay potential must be examined on a link-by-link basis. For example, if a failure affects link i, the following relationships apply:

$$\overline{D}_i = \overline{TTR}_i \tag{B-44}$$

where: $\overline{D}_i$ = Average duration of delay for delayed passengers

$\overline{TTR}_i$ = Average downtime for ith failure.

The probability of delay can be estimated by the following expression:

$$Pr_d = \frac{\sum_i PD_i \lambda_i}{DR} \tag{B-45}$$

where: Pr_d = Probability of being delayed on average trip to partial system failure

PD_i = Number of passengers delayed due to ith failure

λ_i = Failure rate for ith failure

DR = Passenger demand rate.

The expected delay can be estimated by the following relationship:

$$ED = \frac{\sum_i PD_i \lambda_i \overline{TTR}_i}{DR} \tag{B-46}$$

where: ED = Average delay expected on average trip due to partial system failure.

(Other variables defined as before.)

The above relationships deal with any delay type. For determination of stoppage potential, the procedure is identical except that only failures which induce en-route stoppage are considered.

PART 4

THE DEVELOPMENT OF MEASURES OF SERVICE AVAILABILITY

TASK 5. SERVICE AVAILABILITY WORKSHOP

Contract No. DOT-TSC-1283

to

DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER

by

R. D. Leis

BATTELLE
Columbus Laboratories
505 King Avenue
Columbus OH 43201

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	4-1
2.	WORKSHOP OBJECTIVES	4-2
3.	SUMMARY OF WORKSHOP ORGANIZATION	4-3
	Workshop Participants	4-3
	Workshop Format	4-3
	Group Discussions	4-4
	Group Reports	4-4
4.	WORKSHOP CONCLUSIONS	4-4
	Passenger Delay Measures	4-4
	Excess Capacity	4-6
	Buyer-Seller Dialogue	4-6
	Compliance Testing	4-6
	System Specifications	4-6
	Assessment of Research Results	4-8
	Guideline Document	4-8
	Service Availability Vis-a-Vis Other Considerations	4-8

APPENDIX A

SERVICE AVAILABILITY WORKSHOP - LIST OF PARTICIPANTS 4-11

APPENDIX B

SERVICE AVAILABILITY WORKSHOP - PROGRAM AGENDA 4-15

APPENDIX C

GRAPHIC MATERIALS 4-16

APPENDIX D

SERVICE AVAILABILITY WORKSHOP - WORKSHOP GROUP FORMATION 4-17

APPENDIX E

SERVICE AVAILABILITY WORKSHOP - QUESTIONS FOR WORKSHOP GROUPS 4-18

APPENDIX F

SERVICE AVAILABILITY WORKSHOP - WORKSHOP GROUP REPORTS 4-21

LIST OF TABLES

Table 4-1. Definition of Service Availability Measures for System Planning and Design 4-5

LIST OF FIGURES

Figure 4-1. Conceptual Service Availability Control Process 4-7

1. INTRODUCTION

At present there exists no standard approach, no standard terminology, and no standard methodological framework for establishing transportation system performance goals and controlling system design and operational parameters pursuant to these goals. This study (which is part of UMTA's Automatic Guideway Transit Technology program) is aimed at developing a set of measures for service availability (for AGT systems) which will be meaningful, understandable, and acceptable to transit operators, suppliers, and interested Government agencies. In review, the major objective of this study was twofold.

(1) To define service availability measures that suitably characterize transit passenger delays

(2) To establish methods for translating passenger service availability measures into system hardware and operational specifications.

Service availability is defined in a generic sense as a measure of the impingement of equipment failures on the operation of a transit system as perceived by the system users and operators.

Task 1 of this study consisted of an in-depth review of existing literature dealing directly or indirectly with service availability. Specifically sought were definitions, use, methods of measurement, models, and concepts as treated in the literature. The results of this effort were reported in Reference 1. Task 2 carried this information-gathering activity to the transit industry to gain the benefit of its members' experience in the use of service availability measures. It was desired to obtain "real-world" insight into the following:

Service availability concepts/definitions

Use of service availability measures in various phases of a transit system's life cycle

Factors influencing service availability and its use

Characteristics of a "good" measure of service availability

Criteria by which alternative measures can be evaluated.

The results of this activity were reported in Reference 2.

The original program schedule called for the subsequent performance of a task of selection of appropriate service availability measures and a task to develop and demonstrate the methodology for utilizing those measures. An important conclusion of Task 2 was that such separation cannot be made. Perhaps the most important criterion for a good measure is the existence of a simple, understandable, and usable methodology for its use. Hence, these tasks (3 and 4) were combined and their results are reported in Reference 3. The report, in focusing on passenger delays, as opposed to equipment delays, sets forth a refined definition of the general statement of service availability, i.e., the impingement of system failures on passengers. The report goes on to show that it is not feasible to obtain general analytic formulations of the dynamic relationships between passenger demand patterns and system physical and operating characteristics applicable to all possible system configurations and service patterns. Rather, it is necessary to treat each system separately and to examine the unique properties which characterize passenger delay dynamics.

Task 5 of the project, Service Availability Workshop, was designed to provide a forum for the preliminary results of the project and to provide an opportunity for those concerned with questions of service availability to react to these results.

2. WORKSHOP OBJECTIVES

The primary objective of the workshop was to have an in-depth exchange between the project staff and industry and Government representatives in order to

(1) Communicate to the transit industry the results of this study of service availability

(2) To obtain a critical review by the industry as to the potential use and limitations of the research.

3. SUMMARY OF WORKSHOP ORGANIZATION

Workshop Participants

Participation in the service availability workshop was by invitation. It was intended that the participants represent a broad spectrum of industry, including operators, designers, and equipment suppliers, as well as Government representatives. To this end, invitations were extended to suppliers, to eleven cities actively planning for DPM (Downtown People Mover) systems, to a number of transit system operators and consultants, and to Federal officials involved in transit system planning and implementation activities. The list of participants, Appendix A, evidences that a broad spectrum of expertise was in attendance.

Workshop Format

The agenda that was distributed to all participants is included as Appendix B. Following a general presentation of the objectives of the research, an overview of the results to date, and a summary of the service availability methodology developed, the participants took part in working group sessions in which each group member had the opportunity to ask questions and to present his views on the results of the research. The basic source document for the workshop (see Appendix C) included a summary of the service availability methodology developed in the project and essential background material for the workshop group discussions.

To the extent possible, each workshop group was composed of representatives from the Federal Government, state and local governments, system suppliers, system operators, and research and consulting firms (see Appendix D). This was designed to encourage a broad range of discussion and interchange of ideas and opinions within each group.

Following the workshop group sessions, all participants met together for a final plenary session for summary statements by each of the group chairmen and a general discussion of accomplishments of the research and the workshop tasks.

Group Discussions

In order to stimulate and direct the workshop group discussions, a list of questions was distributed to each participant (see Appendix E). The questions helped to focus the group discussions and provided the basis for final discussion and concluding remarks when the entire group met in the plenary session. Within each workshop group, a representative from UMTA or TSC acted chairman pro tem and a member of the BCL staff acted as a recording secretary for the group. Mr. R. D. Leis, BCL program manager and Mr. C. W. Watt, TSC program monitor, circulated through the groups to answer questions and to provide additional quidance relative to the BCL service availability approach.

Group Reports

Each group prepared a brief statement in response to the questionnaire; these statements are included in Appendix F.

4. WORKSHOP CONCLUSIONS

The purpose of this section is to summarize the key conclusions of the workshop as a result of the workshop group discussions and the final plenary session. These conclusions relate to a range of topics, including the technical approach and its limitations, the use and usefulness of the results, areas of continuing concern, and suggestions for final presentation. The conclusions presented below draw from the group reports and the discussion of the final plenary session. It is not intended that they exhibit a rank ordering of importance or consensus.

Passenger Delay Measures

It seems apparent that passenger delay measures (SAM 1 - see Table 4-1) are desirable; problems arise, however, when one tries to measure passenger delays in day-to-day transit operation or to use SAM 1 as a system specification. Moreover, a given SAM 1 does not uniquely determine a SAM 2. On the other hand, a given SAM 2 can yield a SAM 1 for a given system. Thus, from the point of view of system specification, SAM 2 may be much more useful if there is an accepted methodological approach for determining a set of SAM 1's for comparison purposes.

TABLE 4-1. DEFINITION OF SERVICE AVAILABILITY MEASURES FOR SYSTEM PLANNING AND DESIGN

Service Availability Measure	General Description	Examples of Measures
SAM 0	Generalized passenger perceptions of delays associated with system failures	• Perceptions of frequency and duration of delays
SAM 1	Acceptable values of passenger delay parameters	• Probability of a delayed trip • Expected delay time given a delayed trip • Maximum limit on cumulative passenger delay per unit time period (day, month, year)
SAM 2	Acceptable values of system/ equipment design parameters	• System excess capacity • Mean time between failures (MTBF) by failure mode • Mean time to repair (MTTR) by failure mode)

Source: "The Development of Measures of Service Availability", Attachment 3 of this report; presented to all workshop participants.

Excess Capacity

It was concluded that excess capacity is an important system characteristic, worth even a substantial additional cost; however, it is not yet clear how excess capacity can be defined, for example, for a planned DPM system in view of variations in O-D patterns, in queue dissipation rates, etc. TTR (time to restore) may be more important than excess capacity if it is assumed that en-route delays associated with stopped vehicles are more serious than station delays.

Buyer-Seller Dialogue

The "SAM 1-f_2-SAM 2 process (see Figure 4-1) implies a major dialogue between buyer and seller. Changes in existing procurement regulations may be necessary to allow for it; alternatively, it may be possible to express system availability specifications in the form of a range of numbers, or a threshold value. Acting within the existing procurement procedure, it may be desirable, if possible, to build problem solving/negotiating provisions into the contract package.

Compliance Testing

SAM 2 is significantly easier for suppliers to deal with in compliance (acceptance) testing than SAM 1; suppliers can be held accountable for performance they can control, i.e., MTBF, MTTR, throughput capacity, etc. SAM 1-type passenger delay measures do not solely reflect system operating characteristics. Compliance testing should involve a lengthy test period or follow an initial operating period of some length (i.e., six months), if this is possible.

System Specifications

There appears to be a consensus that most failures and associated delays will be of only minor concern to DPM passengers as long as vehicles are moving. (The delay situation is further confused because station delays may be perceived as being different than en-route delays.) Passengers on the relatively short DPM systems may be insensitive to station delays of up to 50 percent of

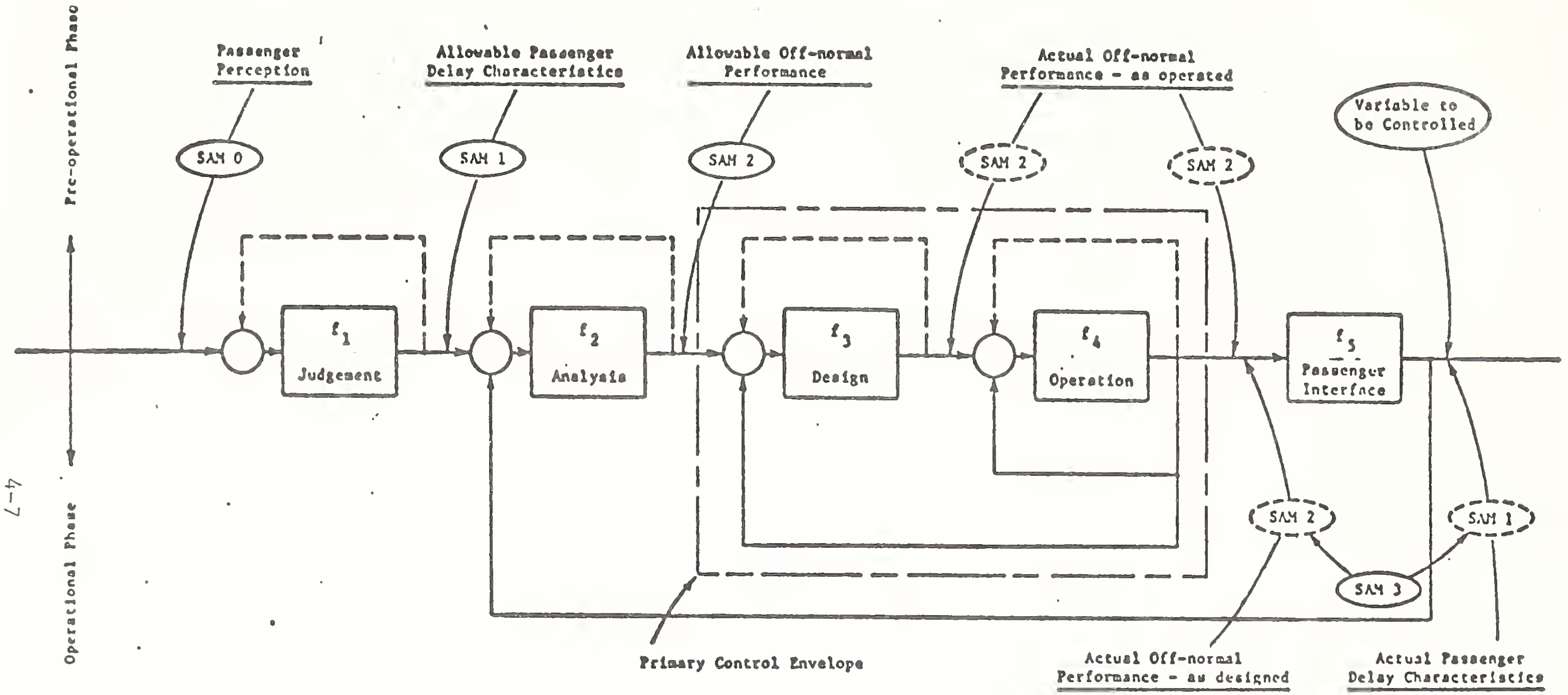


FIGURE 4-1. **CONCEPTUAL SERVICE AVAILABILITY CONTROL PROCESS**

Source: (Same as Table 4-1).

the expected trip time. There was general agreement among suppliers and buyers that specifications such as (a) maximum allowed time to restore service or (b) downtime limits are desirable.

Assessment of Research Results

The research results which have been accomplished thus far were thought to be of high quality. The methodology may be more useful for assessing system performance and workability than for system specification. A real-world evaluation of the f_2-process (translation between SAM 1 and SAM 2) is needed to provide a basis for its ultimate value; the SEA-TAC Satellite Transit System was cited as a potentially good test case. In addition it was noted that a program for evaluating this and other service availability measures would be of value at this time. The reports on Tasks 1 and 2 were also commended; it may be desirable to give them wide distribution.

Guideline Document

The workshop participants were in agreement on the need for a guideline document, or procedure manual, for applying the suggested methodology. The document should provide step-by-step directions for use, as well as examples, so that even those who are not completely conversant in the methodology can use it. This implies that a thorough explanation of what the methodology is intended to do together with its limitations must be included.

Service Availability Vis-a-Vis Other Considerations

The workshop group discussions brought out the need to recognize that service availability is but one of many AGT considerations that must be examined when comparing or evaluating systems. The proper perspective of service availability is conceded to be a difficult, but important task. There is agreement that cities with a DPM system need a method to determine the "goodness" (or "badness") of a system. The problem is one of defining goodness and measuring it in such a way that the performance requirements can be related to passengers' perception of level (quality) of service (LOS). Here, again, complications arise because some characteristics of LOS are related to the physical environment and others to the operating environment; the supplier is responsible for the operating environment, but the difficulty arises in relating the operating

characteristics of the system to the passenger demand in order to provide some level of service.

APPENDIX A

SERVICE AVAILABILITY WORKSHOP

LIST OF PARTICIPANTS

William E. Bell
General Manager
Indianapolis Public Transportation Corporation
P.O. Box 2383
Indianapolis, Indiana 46206 Ph: (317) 635-2100

James D. Boldig
Senior Project Engineer
Transportation Systems Division - General Motors
General Motors Technical Center
Warren, Michigan 48090 Ph: (313) 575-8430

Harvey J. Christiansin
DPM Systems Engineer
Boeing Aerospace Company
Kent Space Center
P.O. Box 3999
Seattle, Washington 98124 Ph: (206) 773-9033

Austin Corbin, Jr.
AIRTRANS Program Manager
Vought Corporation
P.O. Box 5907
Dallas, Texas 75222 Ph: (214) 266-2363

Dennis M. Elliott
Consultant
Dallas-Fort Worth Airport
P.O. Box Drawer DFW
Dallas/Fort Worth Airport, Texas 75261 Ph: (206) 435-5407

Howard D. Evoy
Program Manager AGT Socio-Economic Research Program
Urban Mass Transportation Administration
2100 2nd Street, S.W.
Washington, D.C. 20590 Ph: (214) 574-3062

Roy Field
Transportation Safety Specialist
Urban Technology Development & Deployment
Office of Safety and Product Qualification
2100 2nd Street, S.W.
Washington, D.C. 20590 Ph: (202) 424-9545

Don Gardner
Sr. Electrical Engineer (Project Mgr. for APTA AGTT/AGRT Project)
Southern California RTD
425 S. Main Street
Los Angeles, California 90013 Ph: (213) 972-6437

Robert C. Hoyler
Project Manager
Urban Mass Transportation Administration
2100 2nd Street, S.W.
Washington, D.C. 20590 Ph: (202) 426-4047

Ronald D. Kangas
General Engineer
Transportation Systems Center
Kendall Square
Cambridge, Massachusetts 02142 Ph: (617) 494-2298

John J. Marino
General Engineer
Office of AGT Applications, UTD-60
Urban Mass Transportation Administration
2100 2nd Street, S.W.
Washington, D.C. 20590 Ph: (202) 426-2896

Donald J. Marsh
Technical Engineer
Transportation Systems Division - General Motors
General Motors Technical Center
Warren, Michigan 48090 Ph: (313) 575-8464

Robert E. Murphy
Transportation Project Engineer
Kaizer Engineers
300 Lakeside Drive
Oakland, California 94666 Ph: (415) 271-4059

Robert Jon Pawlak
General Engineer, Safety and System Assurance
Transportation Systems Center
Kendall Square
Cambridge, Massachusetts 02142 Ph: (617) 494-2215

Robert L. Pearson
Project Manager - Transit Systems
Metropolitan Transit Commission
160 East Kellogg
St. Paul, Minnesota 55101 Ph: (612) 221-0939

Mr. W. J. Roesler
Mathematician
Applied Physics Laboratory - Johns Hopkins University
Johns Hopkins Road
Laurel, Maryland 20810 Ph: (301) 953-7100

Mr. Samuel Rudofsky
Manager, Communications & Hardware Section
Michigan Department of State
Highway and Transportation, Bureau of Urban & Public Transportation
425 West Ottawa Street
Lansing, Michigan 48909 Ph: (517) 374-9189

Frederick J. Rutyna
Chief, Automation Branch
Transportation Systems Center
55 Broadway
Cambridge, Massachusetts 02142 Ph: (617) 494-2770

C. W. Watt
General Engineer
Transportation Systems Center
Kendall Square
Cambridge, Massachusetts 02142 Ph: (617) 494-2298

T. D. Whitenack, Jr.
Program Manager-AGRT
Rohr Industries, Inc.
P.O. Box 878
Chula Vista, California 92012 Ph: (714) 575-2037

Dr. William C. Womack
Project Engineer
Otis-TTD
P.O. Box 7293 Park Hill Station
Denver, Colorado 80207 Ph: (303) 343-8780

Mr. Sweson Yang
Principal Planner
Department of Metropolitan Development
Division of Planning and Zoning
2001 City-County Building
Indianapolis, Indiana 46204 Ph: (317) 633-3559

Simon Zweighaft
Manager, Project Development
Metropolitan Dade County - Transportation Administration
44 West Flagler Street, 10th Floor
Miami, Florida 33130 Ph: (305) 579-5323

Walter J. Diewald
Research Engineer
Transportation Systems Section
Battelle-Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201 Ph: (614) 424-5015

C. William Hamilton
Research Leader
Transportation Systems Section
Battelle-Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201 Ph: (614) 424-5131

Rolland D. King
Associate Manager
Transportation Systems Section
Battelle-Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201 Ph: (614) 424-5162

Richard D. Leis
Research Scientist
Transportation Systems Section
Battelle-Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201 Ph: (614) 424-5149

Donald J. Ochsner
Research Scientist
Transportation Systems Section
Battelle-Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201 Ph: (614) 424-4826

APPENDIX B

SERVICE AVAILABILITY WORKSHOP
PROGRAM AGENDA

October 26

8:00	Registration	
9:00	Welcome	R. D. King
9:10	Introductory Remarks	C. W. Watt
9:20	Project Overview	C. W. Hamilton
9:40	Service Availability Control Process	R. D. Leis
10:30	Break	
10:45	Passenger Delays/System Characteristics	R. D. Leis
12:00	Lunch	
1:15	Service Availability Control Methodology	R. D. Leis
2:15	Questions for Clarification	R. D. Leis
2:45	Break	
3:00	Workshop Group Formation	R. D. King
3:10	Workshop Group Discussions	
5:30	Social Hour	

October 27

9:00	Workshop Group Discussions	
12:00	Lunch	
1:15	Workshop Group Reports	R. D. King
2:30	General Discussion and Concluding Remarks	R. D. Leis/ C. W. Watt

APPENDIX C

GRAPHIC MATERIALS

Appendix C included copies of selected figures and Sections 4.0, 5.0, and 6.0 of the Third Interim Report. Since this report is bound in this volume, those sections have not been repeated here.

APPENDIX D

SERVICE AVAILABILITY WORKSHOP

WORKSHOP GROUP FORMATION

1	2	3	4
Marino*	Rutyna*	Kangas*	Pawlak*
Gardner	Boldig	Bell	Bowman
Gunter	Elliott	Christiansin	Corbin
Marsh	Field	Evoy	Hoyler
Whitenack	Rudofsky	Pearson	Murphy
Zweighaft	Womack	Roesler	Yang
Ochsner	*King*	*Hamilton*	*Diewald*

* Group chairman

Note: Leis and Watt will be available to all groups to answer questions that may arise.

APPENDIX E

SERVICE AVAILABILITY WORKSHOP

QUESTIONS FOR WORKSHOP GROUPS

(1) What is the experience of the group relative to service availability measures?

(a) What measures were used?

(b) At what stage of system life-cycle--planning, procurement, operation--was it used?

(c) What were the advantages and disadvantages?

(d) Were the measures effective?

(e) What changes in system hardware and/or operating procedures can be credited to the use of a service availability measure?

(2) How would you rank the importance of service availability relative to other attributes of a transportation system?

(3) How important is abnormal station delay? Is it a significant factor in passenger dissatisfaction?

(4) Do planners/operators know what level of system performance they want relative to passenger delay parameters? Could they make cost/performance tradeoffs?

(5) What problems may be encountered when passenger delay parameters are used in the service availability requirements of the system specification?

(6) How should "foreseen" nonequipment related off-normal performance be handled by the planner?

(7) Should a performance specification for service availability include allowances for uncontrollable "outside" service interruptions such as vandalism, passenger caused interruptions, etc.?

(8) How should "unforeseen" system operational problems be included in the service availability performance specifications? Should the specification writer add some "margin-of-safety"? How does the planner handle this problem?

(9) The service availability measure approach proposed requires knowledge of many aspects of the transportation system--number and location of stations, network configuration, origins/destinations, passenger demand, vehicle size, normal service frequency and line speed, types of service failure modes, frequency for each, and failure recovery approaches for each. How well are these known during the bidding cycle? Can reasonable estimates be made?

(10) Many of today's AGT systems were a first-of-its-kind and at the bidding stage many details of the systems were not yet finalized. How well are the details of the various sytems known today? Can the supplier do the tradeoffs indicated in f_2?

(11) System excess capacity appears to be the important characteristic of transit systems for enabling rapid recovery following a service failure. From your knowledge, is this a correct interpretation?

(12) The service availability measure approach proposes considerable interaction between the buyer and the supplier during the bidding cycle. Is this feasible? What problems are envisioned? What recommendations would you offer?

(13) Is it reasonable to expect agreement to detailed acceptance test procedures during the bidding phase? Will the supplier know his system well enough? Will the buyer know how he plans to operate well enough?

(14) What tradeoff is a supplier willing to make between

(a) The length (and, hence, cost) of equipment acceptance testing, and

(b) The risk of falsely rejecting the equipment?

(15) To obtain confidence in the test results, the acceptance test period may become long, such as a year, is it possible for the operation and/or maintenance of the system to be performed by other than supplier personnel during that period?

(16) Are the report examples used to explain the service availability control process understandable? If not, what areas are not clear?

(17) Is the process proposed workable in the "real world"? What are the anticipated limitations/pitfalls? How could they be avoided?

APPENDIX F

SERVICE AVAILABILITY WORKSHOP

WORKSHOP GROUP REPORTS

Group 1 Marino (UMTA - Chairman)
Gardner (L. A. RTD - APTA)
Marsh (G. M.)
Whitenack (Rohr)
Zweighaft (Miami DPM)
Ochsner (BCL)

(1) Don Marsh is familiar with the Westinghouse system installation and Ochsner is familiar with the AIRTRANS availability requirements. Both systems have been implemented against classical reliability (MTBF, MTTR for equipment) measures and were only slightly effective. The requirements for Westinghouse included some incentives for good performance; AIRTRANS did not. The definition of "failure" and "system recovery" were hang ups in specification interpretation.

(2) Service availability is a very important parameter. There is a "threshold" availability above which the cost effectiveness is questionable. Above this "threshold" calue, cost tradeoffs should be made. Safety, of course, must remain an independent uncompromising parameter of a transportation system.

(3) Concensus was that "station" waits are more acceptable than "vehicle" waits (stalled vehicles). An exception is Morgantown PRT where open stations are very uncomfortable in foul weather. Station waits are of medium importance but should be handled by proper announcements to stranded passengers. The ultimate decision of relative importance is system specific depending on types of stations, station doors, climate, etc.

(4) It was agreed that such values could be selected by planners/consultants. Cost tradeoffs would be difficult due to system design parameters not being available until after receipt of proposals. Summary was that it is unlikely a very good cost tradeoff could be made.

(5) Concensus was that if they are included in the specs, it is also necessary that a very specific explanation of the methodology to be used be included. Also, if sufficient excess capacity is available, and past systems have had excess capacity, there is little need for the extensive passenger delay studies.

(6) Nonequipment-related problems are not the supplier's responsibility. The owner and planner must be educated to expect those problems and make preparations (money, strategies) to handle them.

(7) Only as a margin of safety in the specifications. This margin of safety may best be handled as excess capacity. TTR of the system must be carefully planned to handle "outside" service interruptions as well as equipment failures.

(8) Same as (6) and (7).

(9) The supplier knows or will develop them all during the proposal preparation but the planner will not know them all prior to the release of the RFP. Concensus was that is is very unlikely that all of the needed information could be learned prior to release of the RFP.

(10) Major suppliers (Vought, WEC, Boeing) could do this now based on data from current operating systems. Other suppliers are questionable.

(11) Yes. It was even suggested that perhaps excess capacity purchase would be more cost effective than attempting to calculate the passenger delays due to failures. There is a risk of additional O&M cost and additional failures of the "extra" vehicles may penalize operation if the extra vehicles are not utilized wisely.

(12) This is not feasible due to system-specific design for each application and suppliers desire to keep some "aces" in the hold for the proposals. It may be possible to use a range of numbers and include the passenger delay methodology which will be used to evaluate proposals.

(13) Yes, Yes, Yes. There may be some disagreement between UMTA and the cities, however, on how much testing "proof" is needed for new concepts.

(14) It is anticipated that suppliers are not interested in very much risk on a "system" basis. It is unfair to expect suppliers to work on system construction for years and have a risk of total rejection or very costly redesign because of a "test". System testing should be a long process with acceptance a normal conclusion when a certain service level is proven. Then system improvement will continue to achieve a higher "level" of service.

(15) No, supplier personnel must control during acceptance period. The owner should have aprallel personnel available to witness and acknowledge conformance with criteria for acceptance. A training period must be allowed after acceptance for owners O&M personnel.

(16) No. There appears to be too many examples. One or two examples with detailed step-by-step explanations should be given.

(17) Concensus was that the process is <u>not</u> workable in the real world. The primary problem is the lack of system-specific information from the suppliers. It does appear to be a good evaluation tool for the buyer to use during proposal evaluation and contractor selection.

<u>Group 2</u> Rutyna (UMTA - Chairman)
Bolding (G. M.)
Elliott (Consultant - St. Louis DPM)
Field (UMTA)
Rudofsky (State of Michigan)
Womak (Otis-TTD)
King (BCL)

(1) Group concluded there are 4 or 5 SAMs available. The group could not say whether or not they were effective. There was a difference of opinion within the group as to whether passenger-oriented specs or system hardware specs was the way to go.

(2) From a passenger's viewpoint the group felt that safety was most important and service availability was second. Safety is necessary to get insurance.

(3) The group felt that all delays are important, but that en-route stoppages are more important since the passenger has fewer options.

(4) There is no history for the new systems to draw upon, therefore, the answers are no and no.

(5) Might induce some additional costs if unrealistic. The interactive process proposed may be prohibited by the procurement laws (state and local).

(6) Be cognizant of it. Ask for options from supplier as how to handle the foreseen nonequipment-related off-normal performance.

(7) No.

(8) Unforeseen system operational problems should not be included in the service availability performance specifications. Should allow the bidder to suggest what he foresees as problems with suggested options.

(9) The person preparing the specs generally knows the network configuration and demand and probably isn't privy to the system hardware capabilities to the level desired. The supplier probably has reasonable estimates. What does the buyer do under these conditions when he has to deal with several supplier systems? There was group concensus that the legal constraints would prohibit the open interaction during the bidding period after RFP. Perhaps pre-RFP discussions could be held.

(10) If the supplier can't do it, nobody can. It would be nice to have it a joint effort. There is general feeling that system hardware characteristics are reasonably known.

(11) It is important. The time to restore is more important. Clearing the people in stalled vehicles is more important than clearing queues in stations.

(12) It is feasible. It would have to be some kind of prebidders conference. The bidding costs would be higher. Considerations should be given to influence legislative changes which would prevent proposed open communications.

(13) The test plan can be agreed to; the detailed acceptance test procedures probably cannot be agreed to during the bid cycle. Get the legal people in early relative to acceptance conditions before the test plan is considered. Some provisions should be included in the specs to allow for the "maturing" of the system in its reliability performance. There was some concern with the "maturing" concept in that the system must work well when it is open to revenue service. Perhaps some minimal level of performance can be established with a further requirement for improvement over some period of time.

(14) The importance of the acceptance test should not be underestimated.

(15) The group thought it was possible. This is particularly true with the training of buyer maintenance personnel. It probably is necessary to have the maintenance under the control of the supplier until acceptance is achieved.

(16) "The elevator business has its ups and downs."

(17) As a process it probably is workable. It probably would require a computer to make it practical. A viewpoint was offered that the buyer must, at a minimum, be able to review the supplier's analysis so that reasonable evaluations can be made.

Group 3 Kangas (UMTA - Chairman)
Bell (Indianapolis Pyblic Transit Corp.)
Christiansin (Boeing)
Pearson (St. Paul Metro. Transit Commission)
Roesler (APL - Johns Hopkins)
Hamilton (BCL)

(1) The service availability methodology may provide a useful communications tool for making tradeoffs between service availability versus cost. On the other hand, the extensive buyer-supplier communication may be too time consuming. Some of the problems with service availability include defining the operational aspects of improving availability, and the fact that different actors control different variables. Compliance testing of service availability requirements requires a demonstration of full fleet operation even if demand does not warrant it; in service compliance testing may not work for a DPM.

(2) Capital, operating and maintenance cost are extremely important. The buyer, rather than the Federal Government should specify the service availability requirements.

(3) DPM station-size limitations may impact these considerations; it may be that en-route delays (for stopped vehicles) should be weighted more heavily than station delays, e.g., Japan/Okinawa. The bill package must define failure.

(4) Examples were given of the Indianapolis bus system (schedule adherence is defined by being not more than one minute ahead of schedule or not more than three minutes behind) and Minneapolis/St. Paul (zero minutes early and not more than five minutes late). Bus adherence has little to do with DPM systems; headway maintenance is more important for a DPM. One operator prefers longer headway and larger vehicles to simplify meeting performance; it was felt that suppliers can make necessary cost/performance tradeoffs.

(5) There was operator and supplier concern with dealing with passenger delays; vehicle headways are controllable but the high variance in PAX demand undermines the utility of using PAX delay statistics. There is a need to put more attention on SAM 2 relative to operations than on SAM 2 relative to vehicle reliability

(6)
(7)
(8) Manufacturers prefer to design to vehicle-related parameters; they can handle specifications regarding maximum allowable time to restore service and/or a downtime limit on the system. The guideline document may address some of these issues.

(9) Information generally known (Minneapolis/St. Paul); don't know reliability--recorded in different ways.

(10) Yes, f_2 can be done. Analysis methodology should be tested at existing site, i.e., Busch Gardens.

(11) Excess capacity is important in theory but, in practice, queues may not build up in DPM resulting in lost patronage. Excess capacity is viewed as capability to handle higher than anticipated demands.

(12) Pre-RFP communication is difficult and post-RFP communication is procedurally restrictive.

(13) Absolutely required

QA plan is an important part of bid specification

Can prepare standout acceptance tests applicable for testing all bidders systems.

(14) Supplier wants a low-risk test

Would run lengths test to provide confidence in results.

(15) The majority of test personnel must be under control of person responsible for outcome (the supplier) otherwise he cannot guarantee results. When does warranty period start; this supplier is paid for system (apart from incentive payments)?

(16) The presentation of results needs to be improved although a good job was done. There is a need for a step-by-step cookbook to explain methodology.

(17) Useful as design tool, possibly a buyers tool to use before RFP.

Group 4 Pawlak (UMTA - Chairman)
Corbin (Vought)
Hoyler (UMTA)
Murphy (Kaiser Engineers - St. Paul DPM)
Yang (Indianapolis Metro. Planning)
Diewald (BCL)

(1) The primary experience of the group was related to the AIRTRANS system. Mr. Corbin of Vought presented us with some information on the operating rules relative to system availability and a compilation of the weekly service availability achievements for 1977 to date (see Attachments 1 and 2).

(2) Service availability is one of many system attributes of importance in describing and assessing a transportation system.

(3) A lot of other factors are probably important in user evaluation of a system, some fixed and some variable; it could be very important if it goes beyond some reasonable threshold.

(4) The answer is no; as yet system performance is not quantifiable as a single factor.

(5) No answer here except that there must be a balance with other system requirements.

(6),(7), & (8) Not really dealt with; too specific for this group. The questions had not been really addressed before by the participants (except Corbin).

(9) The general feeling was that the specifications should be very inclusive and, therefore, as tight as possible. There is a need for continuing interaction and communication with potential suppliers in the specification writing phase.

(10) The general concensus was that although much is known about existing systems, much more data need to be collected to allow for meaningful comparisons.

(11) TTR was deemed to be more important; again the cost trade-offs involved must be carefully analyzed.

(12) Buyer-supplier interaction is essential but the difficulties (well-known) are great; UMTA should do everything it can to promote a bidding cycle that will maximize the interaction.

(13) Specification requirements should be directly related to an acceptance test procedure.

(14) Again, there is great need for buyer-supplier interaction here because no supplier will allow himself to get into a situation where he would have to remove an unacceptable system.

(15) It would not seem to be desirable from a supplier standpoint unless adequate agreements are drawn up.

(16) A more extensive system example would greatly assist in displaying the process.

(17) It was a considered workable process but there was not enough understanding to allow for a clear statement of the limitations/pitfalls.

110 copies

U. S. DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER
KENDALL SQUARE, CAMBRIDGE, MA. 02142

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

U.S.MAIL

POSTAGE AND FEES PAID
U. S. DEPARTMENT OF TRANSPORTATION
513